

11-29-00

SEQ/A



PATENT APPLICATION

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Docket No: 01017/35966A



PATENT APPLICATION TRANSMITTAL UNDER 37 C.F.R. 1.53

Box Patent Application
Commissioner for Patents
Washington, D.C. 20231

Sir:

Transmitted herewith for filing is the patent application of

Inventor(s): Han et al.

Title: The Human E3 α Ubiquitin Ligase Family

1. Type of Application

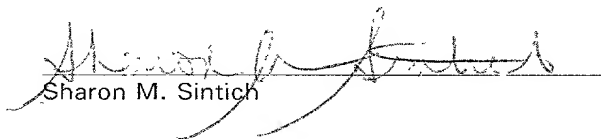
- ☒ This is a new application for a
- ☒ utility patent.
- ☐ design patent.
- ☐ This is a continuation-in-part application of prior application no.

2. Application Papers Enclosed

- 1 Title Page
- 106 Pages of Specification (excluding Claims, Abstract, Drawings & Sequence Listing)
- 11 Page(s) of Claims
- 1 Page(s) of Abstract
- 16 Sheet(s) of Drawings (Figs. 1 to 12)
- ☐ Formal
- ☒ Informal
- 57 Page(s) of Sequence Listing

CERTIFICATION UNDER 37 CFR 1.10

I hereby certify that this Patent Application Transmittal and the documents referred to as enclosed therewith are being deposited with the United States Postal Service on **November 28, 2000**, in an envelope addressed to the Commissioner for Patents, Washington, D.C. 20231 utilizing the "Express Mail Post Office to Addressee" service of the United States Postal Service under Mailing Label No. EM578444884US.


 Sharon M. Sintich

3. Declaration or Oath

- ☐ Enclosed
 - ☐ Executed by (check all applicable boxes)
 - ☐ Inventor(s)
 - ☐ Legal representative of inventor(s)
(37 CFR 1.42 or 1.43)
 - ☐ Joint inventor or person showing a proprietary interest on behalf of inventor who refused to sign or cannot be reached
 - ☐ The petition required by 37 CFR 1.47 and the statement required by 37 CFR 1.47 are enclosed.
See Item 5D below for fee.
- ☒ Not enclosed - the undersigned attorney or agent is authorized to file this application on behalf of the applicant(s). An executed declaration will follow.

4. Small Entity Status

- ☐ Applicant claims small entity status. See 37 CFR 1.27.
- ☐ A small entity statement is(are) attached.

5. Additional Papers Enclosed

- ☐ Preliminary Amendment
- ☐ Information Disclosure Statement
- ☐ Declaration of Biological Deposit
- ☒ Computer readable copy of sequence listing containing nucleotide and/or amino acid sequence
- ☐ Microfiche computer program
- ☐ Associate Power of Attorney
- ☐ Verified translation of a non-English patent application
- ☐ An assignment of the invention
- ☒ Return receipt postcard
- ☒ Other--Statement of Sequence Listing

6. **Priority Applications Under 35 USC 119**

Certified copies of applications from which priority under 35 USC 119 is claimed are listed below and

☐ are attached.

☒ will follow.

COUNTRY	APPLICATION NO.	FILED
US	60/187,911	March 8, 2000

7. **Filing Fee Calculation (37 CFR 1.16)**

A. ☒ **Utility Application**

CLAIMS AS FILED - INCLUDING PRELIMINARY AMENDMENT (IF ANY)						
			SMALL ENTITY		OTHER THAN A SMALL ENTITY	
	NO. FILED	NO. EXTRA	RATE	FEE	RATE	FEE
BASIC FEE				\$355.00		\$710.00
TOTAL	97 -20	= 77	X 9 =	\$	X 18 =	\$1,386.00
INDEP.	15 - 3	= 12	X 40 =	\$	X 80 =	\$960.00
<input checked="" type="checkbox"/> First Presentation of Multiple Dependent Claim			+ 135 =	\$	+ 270 =	\$270.00
Filing Fee:				\$	OR	\$3,326 00

B. ☐ **Design Application (\$160.00/\$320.00)** Filing Fee: \$ _____

C. ☐ **Plant Application (\$245.00/\$490.00)** Filing Fee: \$ _____

D. **Other Fees**

☐ Recording Assignment [Fee -- \$40.00 per assignment] \$ _____

☐ Petition fee for filing by other than all the inventors or person on behalf of the inventor where inventor refused to sign or cannot be reached [Fee -- \$130.00] \$ _____

☐ Other \$ _____

Total Fees Enclosed \$3,326.00

8. Method of Payment of Fees

- ☒ Enclosed check in the amount of: \$3,326.00
- ☐ Charge Deposit Account No. 13-2855 in the amount of: \$ _____
A copy of this Transmittal is enclosed.
- ☐ Not enclosed

9. Deposit Account and Refund Authorization

The Commissioner is hereby authorized to charge any deficiency in the amount enclosed or any additional fees which may be required during the pendency of this application under 37 CFR 1.16 or 37 CFR 1.17 or under other applicable rules (except payment of issue fees), to Deposit Account No. 13-2855. A copy of this Transmittal is enclosed.

Please refund any overpayment to Marshall, O'Toole, Gerstein, Murray & Borun at the address below.

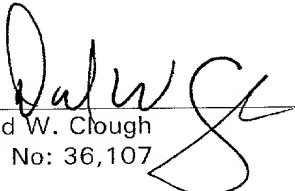
10. Correspondence Address

Customer No.: 04743

Respectfully submitted,

MARSHALL, O'TOOLE, GERSTEIN,
MURRAY & BORUN
6300 Sears Tower
233 South Wacker Drive
Chicago, Illinois 60606-6402
(312) 474-6300
(312) 474-0448 (Telefacsimile)

By:


David W. Clough
Reg. No: 36,107

November 28, 2000

JOINT INVENTORS

"EXPRESS MAIL" mailing label No.

EM578444884US.

Date of Deposit: November 28, 2000

I hereby certify that this paper (or fee) is being deposited with the United States Postal Service "EXPRESS MAIL POST OFFICE TO ADDRESSEE" service under 37 CFR §1.10 on the date indicated above and is addressed to: Commissioner for Patents, Washington, D.C. 20231


Sharon M. Sintich

APPLICATION FOR UNITED STATES LETTERS PATENT

S P E C I F I C A T I O N

TO ALL WHOM IT MAY CONCERN:

Be it known that we, Hui-Quan Han a citizen of the United States of America, residing at 3353 Crossland Street, Thousand Oaks, California, 91362, and Keith Kwak a citizen of the United States of America, residing at 1699 Amarelle Street, Thousand Oaks, California, 91320, have invented new and useful The Human E3 α Ubiquitin Ligase Family, of which the following is a specification.

THE HUMAN E3 α UBIQUITIN LIGASE FAMILY

RELATED APPLICATIONS

This application claims priority under 35 U.S.C. § 119 from U.S.
5 provisional patent application Serial No. 60/187,911 filed March 8, 2000.

FIELD OF THE INVENTION

The present invention includes novel human E3 α ubiquitin ligase
polypeptides (huE3 α I and huE3 α II) and nucleic acid molecules encoding the same. The
10 invention also relates to vectors, host cells, selective binding agents, such as antibodies,
and methods for producing huE3 α polypeptides. Also provided for are methods for the
diagnosis, treatment, amelioration and/or prevention of diseases associated with huE3 α
polypeptides, as well as methods for identifying modulators of huE3 α ligase activity.

BACKGROUND OF THE INVENTION

Technical advances in the identification, cloning, expression and
manipulation of nucleic acid molecules and deciphering of the human genome have
greatly accelerated the discovery of novel therapeutics based upon deciphering of the
human genome. Rapid nucleic acid sequencing techniques can now generate sequence
20 information at unprecedented rates and, coupled with computational analyses, allow the
assembly of overlapping sequences into the partial and entire genomes as well as the
identification of polypeptide-encoding regions. A comparison of a predicted amino acid
sequence against a database compilation of known amino acid sequences can allow one
to determine the extent of homology to previously identified sequences and/or structural
25 landmarks. The cloning and expression of a polypeptide-encoding region of a nucleic
acid molecule provides a polypeptide product for structural and functional analyses. The
manipulation of nucleic acid molecules and encoded polypeptides to create variants and
derivatives thereof may confer advantageous properties on a product for use as a
therapeutic.

30 In spite of significant technical advances in genome research over the past
decade, the potential for the development of novel therapeutics based on the human
genome is still largely unrealized. Many genes encoding potentially beneficial

polypeptide therapeutics, or those encoding polypeptides which may act as “targets” for therapeutic molecules, have still not been identified. In addition, structural and functional analyses of polypeptide products from many human genes have not been undertaken.

Accordingly, it is an object of the invention to identify novel polypeptides and nucleic acid molecules encoding the same which have diagnostic or therapeutic benefit.

Most types of intracellular proteins are degraded through the ubiquitin-proteasome pathway. In this system, proteins are marked for proteosomal degradation by the conjugation of ubiquitin molecules to the protein. Conjugation of the ubiquitin molecule initially involves activation by the E1 enzyme. Upon activation the ubiquitin molecule is transferred to the E2 enzyme which serves as a carrier-protein. The E2 enzyme interacts with a specific E3 ligase family member. The E3 ligase binds to proteins targeted for degradation and catalyzes the transfer of ubiquitin from the E2 carrier enzyme to the target protein. Since the target protein binds to the ligase prior to conjugation, E3 ligase is the rate limiting step for ubiquitin conjugation and determines the specificity of the system. The ubiquitin chain serves as a degradation marker for the 26S proteasome (See Ciechanover, *EMBO J.*, 17: 7151-7160, 1998).

There are only a few known E3 ligases and the sequence homology between them is low. The E3 α family is the main family of intracellular ubiquitin ligases and is involved in N-end rule pathway of protein degradation. The N-end rule states that there is a strong relation between the *in vivo* half-life of a protein and the identity of its N-terminal amino acids. Accordingly, E3 α enzyme binds directly to the primary destabilizing N-terminal amino acid and catalyzes ubiquitin conjugation thereby targeting the protein for degradation. E3 α family members also recognize non-N-end rule substrates (See Ciechanover, *EMBO J.*, 17: 7151-7160, 1998).

The E3 α enzyme family currently consists of intracellular enzymes isolated from rabbit (Reiss and Hershiko, *J. Biol. Chem.* 265: 3685-3690, 1990), mouse (Kwon *et al.*, *Proc. Natl. Acad. Sci., U.S.A* 95: 7898-7903, 1999), yeast (Bartel *et al.*, *EMBO J.*, 9: 3179-3189, 1990) and the *C. elegans* (Wilson *et al.*, *Nature*, 368: 32-38, 1994; Genebank Accession No. U88308) counterparts termed UBR-1. Comparison of these known sequences indicates regions of high similarity regions (I-V) which suggest

the existence of a distinct family. The regions of similarity contain essential residues for the recognition of N-end rule substrates. In region I, the residues Cys-145, Val-146, Gly-173, and Asp-176 are known to be necessary for type-1 substrate binding in yeast and are conserved in the mouse. In regions II and III, residues Asp-318, His-321, and Glu-560 are essential for type-2 substrate binding in yeast and are also conserved in the mouse. In addition, there is a conserved zinc-finger domain in region I and a conserved RING-H2 domain in region IV (Kwon *et al.*, *Proc. Natl. Acad. Sci., U.S.A.*, 95: 7898-7903, 1999).

The full length mouse E3 α cDNA sequence and a partial human E3 α nucleotide sequence (≈ 1 kb) have recently been cloned and characterized as described in US 5,861,312 and Kwon *et al.* (*Proc. Natl. Acad. U.S.A.*, 95: 7898-7908, 1999). The full length mouse E3 α cDNA sequence is 5271 bp in length and encodes a 1757 amino acid polypeptide. The mouse E3 α gene is localized to the central region of chromosome 2 and is highly expressed in skeletal muscle, heart and brain. The partial human E3 α sequence was used to characterize tissue expression and chromosomal localization. This analysis indicated that the human E3 α gene is located on chromosome 15q and exhibits a similar expression pattern as mouse E3 α with high expression in skeletal muscle, heart and brain. As described herein, the present invention discloses two novel, full length, human E3 α sequences (huE3 α I and huE3 α II) and a novel, full length mouse E3 α sequence (muE3 α II). Expression of huE3 α I and huE3 α II mRNA is highly enriched in skeletal muscle tissues. Functionally, huE3 α polypeptides are intracellular enzymes that control protein conjugation and degradation.

Increased proteolysis through the ubiquitin-proteasome pathway has been determined to be a major cause of rapid muscle wasting in many pathological states including but not limited to fasting, metabolic acidosis, muscle denervation, kidney failure, renal cachexia, uremia, diabetes mellitus, sepsis, AIDS wasting syndrome, cancer cachexia, negative nitrogen balance cachexia, burns and Cushing's syndrome (See Mitch and Goldberg, *New England J. Med.*, 335: 1897-1905, 1996). Studies in animal models have shown that muscle wasting disorders are associated with increased ubiquitin content in muscles, increased levels of mRNA transcripts encoding ubiquitin, E2 enzyme and proteasome subunit mRNA, and increased ubiquitin-conjugation to muscle-proteins (See Lecker *et al.*, *J. Nutr.*, 129: 227S-237S, 1999). In this context, the N-end rule

pathway has been shown to play a role in muscle atrophy. E3 α inhibitors, such as dipeptides and methyl ester, reduce the level of ubiquitin conjugation in atrophying rat muscles caused by sepsis, fasting and cancer cachexia (Soloman *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 95: 12602-12607, 1999). These observations indicate that E3 α plays a role in the overall increase in ubiquitination that is associated with and may mediate muscle atrophy in catabolic and other disease states.

Thus, identification of members of the N-end rule protein degradation pathway has led to a better understanding of protein degradation in human cells and the mechanisms of protein degradation in pathological condition which involve muscle atrophy. Identification of the two novel human E3 α ubiquitin ligase genes and polypeptides, as described herein, will further clarify the understanding of these processes and facilitate the development of therapies for pathological conditions which involve abnormal or excessive protein degradation including conditions which involve atrophy of muscle.

SUMMARY OF THE INVENTION

The present invention relates to novel human E3 α nucleic acid molecules and polypeptides encoded by these nucleic acid molecules.

The invention provides isolated nucleic acid molecules comprising or consisting of a nucleotide sequence selected from the group consisting of:

- a) the nucleotide sequence as set forth in SEQ ID NOS: 1 or 3 ;
- b) a nucleotide sequence encoding the polypeptide set forth in SEQ ID NOS: 2 and 4;
- c) a nucleotide sequence which hybridizes under moderate or highly stringent conditions to the compliments of (a) or (b); and
- d) a nucleotide complementary to (a) - (c)

The invention also provides isolated nucleic acid molecules comprising a nucleotide sequence selected from the group consisting of:

- a) a nucleotide sequence encoding a polypeptide that is at least about 70, 75, 80, 85, 90, 95, 96, 97, 98, or 99 percent identical to the polypeptide set forth in SEQ ID NOS: 2 or 4, wherein the polypeptide has an activity of the polypeptide set forth in SEQ

ID NOS: 2 or 4 and the percent identity for these nucleic acid sequences are determined using a computer program selected from the group consisting of GAP, BLASTP, BLASTN, FASTA, BLASTA, BLASTX, BestFit, and the Smith-Waterman algorithm;

b) a nucleotide sequence encoding an allelic variant or splice variant of the nucleotide sequence as set forth in SEQ ID NOS: 1 or 3;

c) the nucleotide sequence of the DNA insert in ATCC Deposit No. PTA-1489 or PTA-1490;

d) a nucleotide sequence of SEQ ID NOS: 1; 3; (a); or (b) encoding a polypeptide fragment of at least about 25 amino acid residues, wherein the polypeptide has an activity of the polypeptide set forth in SEQ ID NOS: 2 or 4;

e) a nucleotide sequence of SEQ ID NOS: 1, 3, or (a)-(c) comprising a fragment of at least about 16 nucleotides;

f) a nucleotide sequence which hybridizes under moderately or highly stringent conditions to the complement of any of (a)-(e); and

g) a nucleotide sequence complementary to any of (a)-(d).

The invention also provides isolated nucleic acid molecules comprising a nucleotide sequence selected from the group consisting of:

a) a nucleotide sequence encoding a polypeptide set forth in SEQ ID NOS: 2 or 4 with at least one conservative amino acid substitution, wherein the polypeptide has an activity of the polypeptide set forth in SEQ ID NOS: 2 or 4;

b) a nucleotide sequence encoding a polypeptide set forth in SEQ ID NOS: 2 or 4 with at least one amino acid insertion, wherein the polypeptide has an activity of the polypeptide set forth in SEQ ID NOS: 2 or 4;

c) a nucleotide sequence encoding a polypeptide set forth in SEQ ID NOS: 2 or 4 with at least one amino acid deletion, wherein the polypeptide has an activity of the polypeptide set forth in SEQ ID NOS: 2 or 4;

d) a nucleotide sequence encoding a polypeptide set forth in SEQ ID NOS: 2 or 4 which has a C- and/or N- terminal truncation, wherein the polypeptide has an activity of the polypeptide set forth in SEQ ID NOS: 2 or 4;

e) a nucleotide sequence encoding a polypeptide set forth in SEQ ID NOS: 2 or 4 with at least one modification selected from the group consisting of amino acid

substitutions, amino acid insertions, amino acid deletions, C-terminal truncation, and N-terminal truncation, wherein the polypeptide has an activity of the polypeptide set forth in SEQ ID NOS: 2 or 4;

- f) a nucleotide sequence of (a)-(e) comprising a fragment of at least about 16 nucleotides;
- g) a nucleotide sequence which hybridizes under moderately or highly stringent conditions to the complement of any of (a)-(f); and
- h) a nucleotide sequence complementary to any of (a)-(e).

The invention also provides isolated polypeptides comprising the amino acid sequence selected from the group consisting of:

- a) the amino acid sequence as set forth in SEQ ID NOS: 2 or 4;
- b) the mature amino acid sequence as set forth in SEQ ID NOS: 2 or 4 comprising a mature amino terminus at residues 1, and optionally further comprising an amino terminal methionine;
- c) an amino acid sequence that is at least about 70, 75, 80, 85, 90, 95, 96, 97, 98, or 99 percent identical to the amino acid sequence of the polypeptide of SEQ ID NOS: 2 or 4 wherein the polypeptide has an activity of the polypeptide set forth in SEQ ID NOS: 2 or 4 and the percent identity for these amino acid sequences are determined using a computer program selected from the group consisting of GAP, BLASTP, BLASTN, FASTA, BLASTA, BLASTX, BestFit, and the Smith-Waterman algorithm.
- d) a fragment of the amino acid sequence set forth in SEQ ID NOS: 2 or 4 comprising at least about 25, 50, 75, 100, or greater than 100 amino acid residues, wherein the fragment has an activity of the polypeptide set forth in SEQ ID NOS: 2 or 4;
- e) the amino acid sequence encoded by the DNA insert of ATCC Deposit No. PTA-1489 or PTA-1490;
- f) an amino acid sequence for an ortholog of SEQ ID NOS: 2 or 4; including the murine ortholog set out as SEQ ID NO: 6.
- g) an allelic variant or splice variant of (a), (b), (e) or (f);

The present invention also provides isolated polypeptides comprising the amino acid sequence selected from the group consisting of:

a) the amino acid sequence set forth in SEQ ID NOS: 2 or 4 with at least one conservative amino acid substitution, wherein the polypeptide has an activity of the polypeptide set forth in SEQ ID NOS: 2 or 4;

b) the amino acid sequence as set forth in SEQ ID NOS: 2 or 4 with at least one amino acid insertion, wherein the polypeptide has an activity of the polypeptide set forth in SEQ ID NOS: 2 or 4;

c) the amino acid sequence as set forth in SEQ ID NOS: 2 or 4 with at least one amino acid deletion, wherein the polypeptide has an activity of the polypeptide set forth in SEQ ID NOS: 2 or 4;

d) the amino acid sequence as set forth in SEQ ID NOS: 2 or 4 which has a C- and/or N- terminal truncation, wherein the polypeptide has an activity of the polypeptide set forth in SEQ ID NOS: 2 or 4; and

e) the amino acid sequence as set forth in SEQ ID NOS: 2 or 4, with at least one modification selected from the group consisting of amino acid substitutions, amino acid insertions, amino acid deletions, C-terminal truncation, and N-terminal truncation, wherein the polypeptide has an activity of the polypeptide set forth in SEQ ID NOS: 2 or 4.

The present invention provides expression vectors comprising the nucleic acid molecules set forth herein, host cells comprising the expression vectors of the invention, and a method of producing a human E3 α polypeptide comprising culturing the host cells and optionally isolating the polypeptide so produced. Another embodiment provides for viral vectors comprising the nucleic acid molecules of the inventions. Further provided is a process for determining whether a compound inhibits huE3 α polypeptide activity or production comprising exposing a host cell expressing huE3 α polypeptide to the compound, and measuring huE3 α polypeptide activity or production in said cell.

A transgenic non-human animal comprising a nucleic acid molecule encoding a huE3 α polypeptide is also encompassed by the invention. The huE3 α nucleic acid molecules are introduced into the animal in a manner that allows expression and

increased levels of a huE3 α polypeptide, which may include increased circulating levels. The transgenic non-human animal is preferably a mammal, and more preferably a rodent, such as a rat or a mouse.

Also provided are derivatives of the huE3 α polypeptides of the present invention, fusion polypeptides comprising the huE3 α polypeptides of the invention, and selective binding agents such as antibodies capable of specifically binding the polypeptides of the invention.

Pharmaceutical compositions comprising the nucleotides, polypeptides, or selective binding agents of the present invention and a carrier, adjuvant, solubilizer, stabilizer, anti-oxidant, or other pharmaceutically acceptable formulation agent are also encompassed by the invention. The pharmaceutical compositions include therapeutically effective amounts of the nucleotides or polypeptides of the present invention, and involve methods of using the polypeptides and nucleic acid molecules.

The huE3 α polypeptides and nucleic acid molecules of the present invention may be used for therapeutic or diagnostic purposes to treat, prevent, and/or detect diseases or disorders, including those recited herein.

Methods of regulating expression and modulating (*i.e.*, increasing or decreasing) levels of a huE3 α polypeptide are also encompassed by the invention. One method comprises administering to an animal a nucleic acid molecule encoding a huE3 α polypeptide. In another method, a nucleic acid molecule comprising elements that regulate or modulate the expression of a huE3 α polypeptide may be administered. Examples of these methods include gene therapy, cell therapy and antisense therapy as further described herein. Further provided is a method of identifying a compound which binds to a huE3 α polypeptide comprising.

A device, comprising a membrane suitable for implantation and host cells expressing a huE3 α polypeptide encapsulated within said membrane, wherein said membrane is permeable to said protein product and impermeable to materials detrimental to said cells is also encompassed by the present invention.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 shows the alignment of the amino acid sequences for huE3 α I, huE3 α II, muE3 α I and muE3 α II (SEQ ID NOS: 2, 4, 15 and 6, respectively).

Figure 2 shows the results of a human multiple tissue Northern blot detecting huE3 α II expression.

Figure 3 shows the results of a human multiple tissue Northern blot detecting huE3 α I expression.

Figure 4 shows that transfection of 293T cells with huE3 α I and huE3II cDNA stimulates the ubiquitination of endogenous proteins and exogenously added α -lactalbumin in cell lysates. The left panel shows the results of gel-shift assays of ubiquitinated proteins. The high molecular weight bands (above 18 kDa for endogenous proteins and above 33 kDa for α -lactalbumin) are identified as “¹²⁵I-Ubiquitin-protein conjugates”. The left panel plots the quantitative measurement of ubiquitinated proteins measured by a PhosphoImager.

Figure 5 shows that transfection of C₂C₁₂ and L6 myotube cells with huE3 α I and huE3II cDNA stimulates the ubiquitination of endogenous proteins cell lysates. The left panel shows the ubiquitinated high molecular weight bands (above 18 kDa for endogenous proteins) as “¹²⁵I-Ubiquitin-protein conjugates”. The left panel plots the quantitative measurement of ubiquitinated proteins measured by a PhosphoImager.

Figure 6 shows the ¹²⁵I-ubiquitin conjugation to endogenous muscle proteins and its sensitivity to selective inhibitors of E3 α in muscle extracts from control and YAH-tumor bearing rats. Gel-shift assays of muscle extracts from control and tumor-bearing rats revealed the ubiquitinated high molecular weight bands (above 18 kDa) denoted as “¹²⁵I-Ubiquitin-protein conjugates”. The left panel is muscle extracts collected 3 days post-implantation and the right panel is muscle extracts collected 5 days post-implantation.

Figure 7 shows the ubiquitin conjugation to ¹²⁵I- α -lactalbumin in extracts from atrophying muscles in YAH-tumor bearing rats as western blots of muscle extracts from control and tumor-bearing rats with the ubiquitinated high molecular weight bands (above 33 kDa) as “¹²⁵I-Lactalbumin-ubiquitin conjugation”. The left panel is muscle extracts collected 3 days post-implantation and the right panel is muscle extracts collected 5 days post-implantation.

Figure 8 shows Northern blot analysis of E3 α I and E3 α II expression in skeletal muscle in YAH-130 experimental cachexia model. The RNA expression

from pair-fed control rats and tumor-bearing rats were compared 3 days (3d) and 5 days (5d) post-implantation.

Figure 9 shows Northern blot analysis of E3 α I and E3 α II expression in skeletal (gastrocnemius) muscle and cardiac muscle in the C26 experimental cacheixia model. The RNA expression from pair-fed control rats and tumor-bearing mice were compared 12 days (12d) and 17 days (17d) post-implantation.

Figure 10 shows induction of E3 α II expression by proinflammatory cytokines TNF α and IL-6 in C₂C₁₂ myotube cultures on Northern blots. The RNA levels of E3 α II (upper panel) and E3 α I (lower panel) were detected 3 or 5 days after treatment with TNF α (left panel) and IL-6 (right panel).

Figure 11 shows that IL-6 treatment causes a time-dependent acceleration of ubiquitination in differentiated C₂C₁₂ cells. This data exhibits the results of a gel-shift assay showing the ubiquitinated high molecular weight bands denoted as “¹²⁵I-ubiquitin protein conjugates”(left panel) and is quantitated by a PhosphoImager in the right panel.

Figure 12 shows that TNF α treatment causes a dose-dependent acceleration of ubiquitination in differentiated C₂C₁₂ cells. This data is displayed as gel-shift assay results with the ubiquitinated high molecular weight bands denoted as “¹²⁵I-ubiquitin protein conjugates” (left panel) and is quantitated by a PhosphoImager in the right panel.

DETAILED DESCRIPTION OF THE INVENTION

The section headings used herein are for organizational purposes only and are not to be construed as limiting the subject matter described therein. All references cited in this application are expressly incorporated by reference herein.

Definitions

The term “huE3 α ” encompasses two novel orthologs of human E3 α ubiquitin ligase described herein including huE3 α I polynucleotide and polypeptide (SEQ ID NOS: 1 and 2, respectively) and huE3 α II polynucleotide and polypeptide (SEQ ID NOS: 3 and 4, respectively).

The term "huE3 α nucleic acid molecule" or "polynucleotide" refers to a nucleic acid molecules including a nucleotide sequence as set forth in SEQ ID NOS: 1 or 3, a nucleotide sequence encoding the polypeptide set forth in SEQ ID NOS: 2 or 4, a nucleotide sequence of the DNA insert in ATCC deposit nos. PTA-1489 or PTA-1490, or nucleic acid molecule related thereto. Related nucleic acid molecules include a nucleotide sequence that is at least about 70 percent identical to the nucleotide sequence as shown in SEQ ID NOS: 1 or 3, or comprise or consist essentially of a nucleotide sequence encoding a polypeptide that is at least about 70 percent identical to the polypeptide set forth in SEQ ID NOS: 2 or 4. In preferred embodiments, these nucleotide sequences are about 75 percent, or about 80 percent, or about 85 percent, or about 90 percent, or about 95, 96, 97, 98, or 99 percent identical to the nucleotide sequence as shown in SEQ ID NOS: 1 or 3, or the nucleotide sequences encode a polypeptide that is about 75 percent, or about 80 percent, or about 85 percent, or about 90 percent, or about 95, 96, 97, 98, or 99 percent identical to the polypeptide sequence as set forth in SEQ ID NOS: 2 or 4.

Related nucleic acid molecules also include fragments of the huE3 α I or hu E3 α II nucleic acid molecules which fragments contain at least about 10 contiguous nucleotides, or about 15, or about 20, or about 25, or about 50, or about 75, or about 100, or greater than about 100 contiguous nucleotides of a huE3 α nucleic acid molecule of SEQ ID NOS: 1 or 3. Related nucleic acid molecules also include fragments of the above huE3 α nucleic acid molecules which encode a polypeptide of at least about 25 amino acid residues, or about 50, or about 75, or about 100, or greater than about 100 amino acid residues of the huE3 α polypeptide of SEQ ID NOS: 2 or 4. Related nucleic acid molecules also include a nucleotide sequence encoding a polypeptide comprising or consisting essentially of a substitution, modification, addition and/or a deletion of one or more amino acid residues compared to the polypeptide set forth in SEQ ID NOS: 2 or 4. In addition, related huE3 α nucleic acid molecules include those molecules which comprise nucleotide sequences which hybridize under moderately or highly stringent conditions as defined herein with the fully complementary sequence of any of the huE3 α nucleic acid molecules of SEQ ID NOS: 1 or 3.

In preferred embodiments, the related nucleic acid molecules comprise sequences which hybridize under moderately or highly stringent conditions with a molecule having a sequence as shown in SEQ ID NOS: 1 or 3, or of a molecule encoding a polypeptide, which polypeptide comprises the sequence as shown in SEQ ID NOS: 2 or 4, or of a nucleic acid fragment as defined herein, or of a nucleic acid fragment encoding a polypeptide as defined herein or the complement of any or the foregoing molecules. It is also understood that related nucleic acid molecules include allelic or splice variants of a huE3 α nucleic acid molecule of SEQ ID NOS: 1 or 3, and include sequences which are complementary to any of the above nucleotide sequences. The related encoded polypeptides possess at least one activity of the polypeptide depicted in SEQ ID NOS: 2 or 4.

The term "isolated nucleic acid molecule" refers to a nucleic acid molecule of the invention that is free from at least one contaminating nucleic acid molecule with which it is naturally associated. Preferably, the isolated nucleic acid molecule of the present invention is substantially free from any other contaminating mammalian nucleic acid molecule(s) which would interfere with its use in polypeptide production or its therapeutic, diagnostic, or preventative use.

A "nucleic acid sequence" or "nucleic acid molecule" as used herein refer to a DNA or RNA sequence. The terms encompasses molecules formed from any of the known base analogs of DNA and RNA such as, but not limited to 4-acetylcytosine, 8-hydroxy-N6-methyladenosine, aziridinyl-cytosine, pseudoisocytosine, 5-(carboxyhydroxymethyl) uracil, 5-fluorouracil, 5-bromouracil, 5-carboxymethylaminomethyl-2-thiouracil, 5-carboxy-methylamino-methyluracil, dihydrouracil, inosine, N6-iso-pentenyladenine, 1-methyladenine, 1-methylpseudouracil, 1-methylguanine, 1-methylinosine, 2,2-dimethyl-guanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-methyladenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarbonyl-methyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid, oxybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-

methyluracil, N-uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid, pseudouracil, queosine, 2-thiocytosine, and 2,6-diaminopurine.

The term “operably linked” is used as recognized in the art to refer to an arrangement of flanking sequences wherein the flanking sequences so described are configured or assembled so as to perform their usual function. Thus, a flanking sequence operably linked to a coding sequence may be capable of effecting the replication, transcription and/or translation of the coding sequence. For example, a coding sequence is operably linked to a promoter when the promoter is capable of directing transcription of that coding sequence. A flanking sequence need not be contiguous with the coding sequence, so long as it functions correctly. Thus, for example, intervening untranslated yet transcribed sequences can be present between a promoter sequence and the coding sequence and the promoter sequence can still be considered “operably linked” to the coding sequence.

The term “pharmaceutically acceptable carrier” or “physiologically acceptable carrier” as used herein refer to one or more formulation materials suitable for accomplishing or enhancing delivery of the huE3 α polypeptide, huE3 α nucleic acid molecule, or huE3 α selective binding agent as a pharmaceutical composition.

The term “allelic variant” refers to one of several possible naturally occurring alternate forms of a gene occupying a given locus on a chromosome of an organism or a population of organisms.

The term “splice variant” refers to a nucleic acid molecule, usually RNA, which is generated by alternative processing of intron sequences in an RNA transcript of huE3 α polypeptide amino acid sequence.

The term “expression vector” refers to a vector which is suitable for transformation of a host cell and contains nucleic acid sequences which direct and/or control the expression of inserted heterologous nucleic acid sequences. Expression includes, but is not limited to, processes such as transcription, translation, and RNA splicing, if introns are present.

The term “vector” is used as recognized in the art to refer to any molecule (*e.g.*, nucleic acid, plasmid, or virus) used to transfer coding information to a host cell.

The term "transformation" as used herein refers to a change in a cell's genetic characteristics, and a cell has been transformed when it has been modified to contain a new DNA. For example, a cell is transformed where it is genetically modified from its native state. Following transfection or transduction, the transforming DNA may recombine with that of the cell by physically integrating into a chromosome of the cell, may be maintained transiently as an episomal element without being replicated, or may replicate independently as a plasmid. A cell is considered to have been stably transformed when the DNA is replicated with the division of the cell.

The term "transfection" is used to refer to the uptake of foreign or exogenous DNA by a cell, and a cell has been "transfected" when the exogenous DNA has been introduced inside the cell membrane. A number of transfection techniques are well known in the art and are disclosed herein. See, for example, Graham *et al.*, *Virology*, 52: 456, 1973; Sambrook *et al.*, *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratories, New York, 1989; Davis *et al.*, *Basic Methods in Molecular Biology*, Elsevier, 1986; and Chu *et al.*, *Gene*, 13: 197, 1981. Such techniques can be used to introduce one or more exogenous DNA moieties into suitable host cells.

The term "transduction" is used to refer to the transfer of genes from one bacterium to another, usually by a phage. "Transduction" also refers to the acquisition and transfer of eukaryotic cellular sequences by retroviruses.

The term "host cell" is used to refer to a cell which has been transformed, or is capable of being transformed, by a vector bearing a selected gene of interest which is then expressed by the cell. The term includes the progeny of the parent cell, whether or not the progeny is identical in morphology or in genetic make-up to the original parent, so long as the selected gene is present.

The term "highly stringent conditions" refers to those conditions that are designed to permit hybridization of DNA strands whose sequences are highly complementary, and to exclude hybridization of significantly mismatched DNAs. Hybridization stringency is principally determined by temperature, ionic strength, and the concentration of denaturing agents such as formamide. Examples of "highly stringent conditions" for hybridization and washing are 0.015 M sodium chloride,

0.0015 M sodium citrate at 65-68°C or 0.015 M sodium chloride, 0.0015M sodium citrate, and 50% formamide at 42°C. See Sambrook, Fritsch & Maniatis, *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, (Cold Spring Harbor, N.Y. 1989); Anderson *et al.*, *Nucleic Acid Hybridisation: A Practical Approach*, Ch. 4, IRL Press Limited (Oxford, England).

More stringent conditions (such as higher temperature, lower ionic strength, higher formamide, or other denaturing agent) may also be used, however, the rate of hybridization will be affected. Other agents may be included in the hybridization and washing buffers for the purpose of reducing non-specific and/or background hybridization. Examples are 0.1% bovine serum albumin, 0.1% polyvinyl-pyrrolidone, 0.1% sodium pyrophosphate, 0.1% sodium dodecylsulfate, NaDodSO₄, or SDS, ficoll, Denhardt's solution, sonicated salmon sperm DNA (or other non-complementary DNA), and dextran sulfate, although another suitable agents can also be used. The concentration and types of these additives can be changed without substantially affecting the stringency of the hybridization conditions. Hybridization experiments are usually carried out at pH 6.8-7.4, however, at typical ionic strength conditions, the rate of hybridization is nearly independent of pH. (See Anderson *et al.*, *Nucleic Acid Hybridisation: a Practical Approach*, Ch. 4, IRL Press Limited (Oxford, England)).

Factors affecting the stability of DNA duplex include base composition, length, and degree of base pair mismatch. Hybridization conditions can be adjusted by one skilled in the art in order to accommodate these variables and allow DNAs of different sequence relatedness to form hybrids. The melting temperature of a perfectly matched DNA duplex can be estimated by the following equation:

$$T_m(^{\circ}\text{C}) = 81.5 + 16.6(\log[\text{Na}^+]) + 0.41(\%G+C) - 600/N - 0.72(\%\text{formamide})$$

where N is the length of the duplex formed, [Na⁺] is the molar concentration of the sodium ion in the hybridization or washing solution, %G+C is the percentage of (guanine+cytosine) bases in the hybrid. For imperfectly matched hybrids, the melting temperature is reduced by approximately 1°C for each 1% mismatch.

The term "moderately stringent conditions" refers to conditions under which a DNA duplex with a greater degree of base pair mismatching than could occur under "highly stringent conditions" is able to form. Examples of typical "moderately

stringent conditions" are 0.01 M sodium chloride, 0.0015 M sodium citrate at 50-65°C or 0.015 M sodium chloride, 0.0015 M sodium citrate, and 20% formamide at 37-50°C. By way of example, a "moderately stringent" condition of 50°C in 0.015 M sodium ion will allow about a 21% mismatch.

It will be appreciated by those skilled in the art that there is no absolute distinction between "highly" and "moderately" stringent conditions. For example, at 0.015M sodium ion (no formamide), the melting temperature of perfectly matched long DNA is about 71°C. With a wash at 65°C (at the same ionic strength), this would allow for approximately a 6% mismatch. To capture more distantly related sequences, one skilled in the art can simply lower the temperature or raise the ionic strength.

A good estimate of the melting temperature in 1 M NaCl* for oligonucleotide probes up to about 20 nt is given by:

$$T_m = 2^{\circ}\text{C per A-T base pair} + 4^{\circ}\text{C per G-C base pair}$$

*The sodium ion concentration in 6x salt sodium citrate (SSC) is 1M. *See Suggs et al., Developmental Biology Using Purified Genes*, p. 683, Brown and Fox (eds.) (1981).

High stringency washing conditions for oligonucleotides are usually at a temperature of 0-5°C below the T_m of the oligonucleotide in 6x SSC, 0.1% SDS for longer nucleotides.

The term "huE3 α polypeptide" refers to a polypeptide comprising the amino acid sequence of huE3 α I or huE3 α II (SEQ ID NOS: 2 or 4, respectively), and related polypeptides having a natural sequence or mutated sequence. Related polypeptides include: allelic variants; splice variants; fragments; derivatives; substitution, deletion, and insertion variants; fusion polypeptides; and orthologs of the huE3 α polypeptides of SEQ ID NOS: 2 or 4, which possess at least one activity of the polypeptide depicted in SEQ ID NOS: 2 or 4. Human E3 α polypeptides may be mature polypeptides, as defined herein, and may or may not have an amino terminal methionine residue, depending on the method by which they are prepared.

The term "huE3 α polypeptide fragment" refers to a polypeptide that comprises less than the full length amino acid sequence of a huE3 α I or huE3 α II polypeptide set forth in SEQ ID NOS: 2 or 4, respectively. Such huE3 α fragments can be 6 amino acids or more in length, and may arise, for example, from a truncation at the

amino terminus (with or without a leader sequence), a truncation at the carboxy terminus, and/or an internal deletion of one or more residues from the amino acid sequence. Human E3 α fragments may result from alternative RNA splicing or from *in vivo* protease activity. Membrane-bound forms of huE3 α are also contemplated by the present invention. In preferred embodiments, truncations and/or deletions comprise about 10 amino acids, or about 20 amino acids, or about 50 amino acids, or about 75 amino acids, or about 100 amino acids, or more than about 100 amino acids. The polypeptide fragments so produced will comprise about 25 contiguous amino acids, or about 50 amino acids, or about 75 amino acids, or about 100 amino acids, or about 150 amino acids, or about 200 amino acids. Such huE3 α polypeptide fragments may optionally comprise an amino terminal methionine residue. It will be appreciated that such fragments can also be used, for example, to generate antibodies to huE3 α polypeptides.

The term "huE3 α polypeptide variants" refers to huE3 α polypeptides which contain one or more amino acid sequence substitutions, deletions, and/or additions as compared to the huE3 α polypeptide amino acid sequence set forth as huE3 α I or huE3 α II (SEQ ID NOS: 2 or 4, respectively). Variants may be naturally occurring or artificially constructed. Such huE3 α polypeptide variants may be prepared from the corresponding nucleic acid molecules encoding said variants, which have a DNA sequence that varies accordingly from the DNA sequences for wild type huE3 α polypeptides as set forth in SEQ ID NOS: 1 or 3. In preferred embodiments, the variants have from 1 to 3, or from 1 to 5, or from 1 to 10, or from 1 to 20, or from 1 to 25, or from 1 to 50, or from 1 to 75, or from 1 to 100, or more than 100 amino acid substitutions, insertions, additions and/or deletions, wherein the substitutions may be conservative, or non-conservative, or any combination thereof.

One skilled in the art will be able to determine suitable variants of the native huE3 α polypeptide using well known techniques. For example, one may predict suitable areas of the molecule that may be changed without destroying biological activity. Also, one skilled in the art will realize that even areas that may be important for biological activity or for structure may be subject to conservative amino acid substitutions without destroying the biological activity or without adversely affecting the polypeptide structure.

For predicting suitable areas of the molecule that may be changed without destroying activity, one skilled in the art may target areas not believed to be important for activity. For example, when similar polypeptides with similar activities from the same species or from other species are known, one skilled in the art may compare the amino acid sequence of huE3 α polypeptide to such similar polypeptides. After making such a comparison, one skilled in the art can determine residues and portions of the molecules that are conserved among similar polypeptides. One skilled in the art would know that changes in areas of the huE3 α molecule that are not conserved would be less likely to adversely affect the biological activity and/or structure of a huE3 α polypeptide. One skilled in the art would also know that, even in relatively conserved regions, one may substitute chemically similar amino acids for the naturally occurring residues while retaining activity (conservative amino acid residue substitutions).

Additionally, one skilled in the art can review structure-function studies identifying residues in similar polypeptides that are important for activity or structure. In view of such a comparison, one skilled in the art can predict the importance of amino acid residues in a huE3 α polypeptide that correspond to amino acid residues that are important for activity or structure in similar polypeptides. One skilled in the art may opt for chemically similar amino acid substitutions for such predicted important amino acid residues of huE3 α polypeptides.

If available, one skilled in the art can also analyze the three-dimensional structure and amino acid sequence in relation to that structure in similar polypeptides. In view of that information, one skilled in the art may predict the alignment of amino acid residues of huE3 α polypeptide with respect to its three dimensional structure. One skilled in the art may choose not to make radical changes to amino acid residues predicted to be on the surface of the protein, since such residues may be involved in important interactions with other molecules.

Additional methods of predicting secondary structure include “threading” (Jones *et al.*, *Current Opin. Struct. Biol.*, 7(3):377-87 (1997); Sippl *et al.*, *Structure*, 4(1):15-9 (1996)), “profile analysis” (Bowie *et al.*, *Science*, 253:164-170 (1991); Gribskov *et al.*, *Meth. Enzym.*, 183:146-159 (1990); Gribskov *et al.*, *Proc. Nat. Acad. Sci.*, 84(13):4355-4358 (1987)), and “evolutionary linkage” (See Home, *supra*, and Brenner, *supra* 1997).

Moreover, one skilled in the art may generate test variants containing a single amino acid substitution at each amino acid residue. The variants could be screened using activity assays described herein. Such variants could be used to gather information about suitable variants. For example, if one discovered that a change to a particular amino acid residue resulted in destroyed, undesirably reduced, or unsuitable activity, variants with such a change would be avoided. In other words, based on information gathered from such routine experiments, one skilled in the art can readily determine the amino acids where further substitutions should be avoided either alone or in combination with other mutations.

In making such changes, the hydropathic index of amino acids may be considered. Each amino acid has been assigned a its hydropathic index on the basis of its hydrophobicity and charge characteristics. They are: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (-0.4); threonine (-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3); proline (-1.6); histidine (-3.2); glutamate (-3.5); glutamine (-3.5); aspartate (-3.5); asparagine (-3.5); lysine (-3.9); and arginine (-4.5).

The importance of the hydropathic amino acid index in conferring interactive biological function on a protein is generally understood in the art (Kyte *et al.*, *J. Mol. Biol.*, 157: 105-131, 1982). It is known that certain amino acids may be substituted for other amino acids having a similar hydropathic index or score and still retain a similar biological activity. In making changes based upon the hydropathic index, the substitution of amino acids whose hydropathic indices are within ± 2 is preferred, those which are within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred.

It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity, particularly where the biologically functionally equivalent protein or peptide thereby created is intended for use in immunological embodiments, as in the present case.

The U.S. Patent No. 4,554,101 states that the greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with its immunogenicity and antigenicity, *i.e.*, with a biological property of the protein. As detailed in U.S. Patent No. 4,554,101, the following hydrophilicity

values have been assigned to amino acid residues: arginine (+3.0); lysine (+3.0); aspartate (+3.0 \pm 1); glutamate (+3.0 \pm 1); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); threonine (-0.4); proline (-0.5 \pm 1); alanine (-0.5); histidine (-0.5); cysteine (-1.0); methionine (-1.3); valine (-1.5); leucine (-1.8); isoleucine (-1.8); tyrosine (-2.3); phenylalanine (-2.5); and tryptophan (-3.4).

In making changes based upon similar hydrophilicity values, the substitution of amino acids whose hydrophilicity values are within ± 2 is preferred, those which are within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred. One may also identify epitopes from primary amino acid sequences on the basis of hydrophilicity. Through the methods disclosed in U.S. Patent No. 4,554,101 one of skill in the art is able to identify epitopes from within a given amino acid sequence. These regions are also referred to as "epitopic core regions".

Numerous scientific publications have been devoted to the prediction of secondary structure, and to the identification of epitopes, from analyses of amino acid sequences. See Chou *et al.*, *Biochemistry*, 13(2): 222-245, 1974; Chou *et al.*, *Biochemistry*, 113(2): 211-222, 1974; Chou *et al.*, *Adv. Enzymol. Relat. Areas Mol. Biol.*, 47: 45-148, 1978; Chou *et al.*, *Ann. Rev. Biochem.*, 47: 251-276 and Chou *et al.*, *Biophys. J.*, 26: 367-384, 1979. Moreover, computer programs are currently available to assist with predicting antigenic portions and epitopic core regions of proteins. Examples include those programs based upon the Jameson-Wolf analysis (Jameson *et al.*, *Comput. Appl. Biosci.*, 4(1): 181-186, 1998 and Wolf *et al.*, *Comput. Appl. Biosci.*, 4(1): 187-191, 1988, the program PepPlot® (Brutlag *et al.*, *CABS*, 6: 237-245 1990, and Weinberger *et al.*, *Science*, 228: 740-742, 1985) and other new programs for protein tertiary structure prediction (Fetrow *et al.*, *Biotechnology*, 11: 479-483 1993).

In preferred embodiments, the variants have from 1 to 3, or from 1 to 5, or from 1 to 10, or from 1 to 15, or from 1 to 20, or from 1 to 25, or from 1 to 50, or from 1 to 75, or from 1 to 100, or more than 100 amino acid substitutions, insertions, additions and/or deletions, wherein the substitutions may be conservative, as described herein, or non-conservative, or any combination thereof. In addition, the variants can have additions of amino acid residues either at the carboxy terminus or at the amino terminus (with or without a leader sequence).

Preferred huE3 α polypeptide variants include glycosylation variants wherein the number and/or type of glycosylation sites has been altered compared to native huE3 α polypeptide. In one embodiment, huE3 α polypeptide variants comprise a greater or a lesser number of N-linked glycosylation sites. An N-linked glycosylation site is characterized by the sequence: Asn-X-Ser or Asn-X- Thr, wherein the amino acid residue designated as X may be any amino acid residue except proline. The substitution(s) of amino acid residues to create this sequence provides a potential new site for the addition of an N-linked carbohydrate chain. Alternatively, substitutions which eliminate this sequence will remove an existing N-linked carbohydrate chain. Also provided is a rearrangement of N-linked carbohydrate chains wherein one or more N-linked glycosylation sites (typically those that are naturally occurring) are eliminated and one or more new N-linked sites are created. Additional preferred huE3 α variants include cysteine variants, wherein one or more cysteine residues are deleted or substituted with another amino acid (*e.g.*, serine). Cysteine variants are useful when huE3 α polypeptides must be refolded into a biologically active conformation such as after the isolation of insoluble inclusion bodies. Cysteine variants generally have fewer cysteine residues than the native protein, and typically have an even number to minimize interactions resulting from unpaired cysteines.

The term "huE3 α fusion polypeptide" refers to a fusion of huE3 α I or huE3 α II polypeptide, fragment, and/or variant thereof, with a heterologous peptide or polypeptide. IN addition, the polypeptide comprising the amino acid sequence of SEQ ID NO: 2 or 4 or huE3 α polypeptide variant may be fused to a homologous polypeptide to form a homodimer or to a heterologous polypeptide to form a heterodimer. Heterologous peptides and polypeptides include, but are not limited to: an epitope to allow for the detection and/or isolation of a huE3 α fusion polypeptide; a transmembrane receptor protein or a portion thereof, such as an extracellular domain, or a transmembrane and intracellular domain; a ligand or a portion thereof which binds to a transmembrane receptor protein; an enzyme or portion thereof which is catalytically active; a polypeptide or peptide which promotes oligomerization, such as a leucine zipper domain; a polypeptide or peptide which increases stability, such as an immunoglobulin constant region, and a polypeptide which has a therapeutic activity different from the huE3 α polypeptide.

In addition, a huE3 α polypeptide may be fused to itself or to a fragment, variant, or derivative thereof. Fusions can be made either at the amino terminus or at the carboxy terminus of a huE3 α polypeptide. Fusions may be direct with no linker or adapter molecule or indirect using through a linker or adapter molecule. A linker or adapter molecule may be one or more amino acid residues, typically from 20 amino acids residues, or up to about 50 amino acid residues. A linker or adapter molecule may also be designed with a cleavage site for a DNA restriction endonuclease or for a protease to allow for the separation of the fused moieties. It will be appreciated that once constructed, the fusion polypeptides can be derivatized according to the methods described herein.

In a further embodiment of the invention, a huE3 α polypeptide, including a fragment, variant, and/or derivative, is fused to an Fc region of human IgG. Antibodies comprise two functionally independent parts, a variable domain known as "Fab", which binds antigens, and a constant domain known as "Fc", which is involved in effector functions such as complement activation and attack by phagocytic cells. An Fc has a long serum half-life, whereas an Fab is short-lived (Capon *et al.*, *Nature*, 337: 525-31, 1989). When constructed together with a therapeutic protein, an Fc domain can provide longer half-life or incorporate such functions as Fc receptor binding, protein A binding, complement fixation and perhaps even placental transfer (Capon *et al.*, *Nature*, 337: 525-31. 1989). Table I summarizes the use of certain Fc fusions known in the art, including materials and methods applicable to the production of fused huE3 α polypeptides.

Table I

Fc Fusion with Therapeutic Proteins

	Form of Fc	Fusion partner	Therapeutic implications	Reference
5	IgG1	N-terminus of CD30-L	Hodgkin's disease; anaplastic lymphoma; T-cell leukemia	U.S. Patent No. 5,480,981
	Murine Fcγ2a	IL-10	anti-inflammatory; transplant rejection	Zheng <i>et al.</i> , <i>J. Immunol.</i> , 154: 5590-600, 1995
	IgG1	TNF receptor	septic shock	Fisher <i>et al.</i> , <i>N. Engl. J. Med.</i> , 334: 1697-1702, 1996; Van Zee <i>et al.</i> , <i>J. Immunol.</i> , 156: 2221-30, 1996
10	IgG, IgA, IgM, or IgE (excluding the first domain)	TNF receptor	inflammation, autoimmune disorders	U.S. Pat. No. 5,808,029, issued September 15, 1998
15	IgG1	CD4 receptor	AIDS	Capon <i>et al.</i> , <i>Nature</i> 337: 525-31, 1989
	IgG1, IgG3	N-terminus of IL-2	anti cancer, antiviral	Harvill <i>et al.</i> , <i>Immunotech.</i> , 1: 95-105 1995
	IgG1	C-terminus of OPG	osteoarthritis; bone density	WO 97/23614, published July 3, 1997
	IgG1	N-terminus of leptin	anti-obesity	PCT/US 97/23183, filed December 11, 1997
20	Human Ig Cγ1	CTLA-4	autoimmune disorders	Linsley, <i>J. Exp. Med.</i> , 174: 561-9, 1991

In one example, all or portion of the human IgG hinge, CH2 and CH3 regions may be fused at either the N-terminus or C-terminus of the huE3α polypeptides using methods known to the skilled artisan. In another example, a portion of a hinge regions and CH2 and CH3 regions may be fused. The resulting huE3α Fc-fusion polypeptide may be purified by use of a Protein A affinity column. Peptides and proteins fused to an Fc region have been found to exhibit a substantially greater half-life *in vivo* than the unfused counterpart. Also, a fusion to an Fc region allows for dimerization/multimerization of the fusion polypeptide. The Fc region may be a naturally occurring Fc region, or may be altered to improve certain qualities, such as therapeutic qualities, circulation time, reduce aggregation, etc.

The term "huE3 α polypeptide derivatives" refers to huE3 α I or huE3 α II polypeptides, fragments, or variants, as defined herein, that have been chemically modified. The derivatives are modified in a manner that is different from naturally occurring huE3 α , polypeptides either in the type or location of the molecules attached to the polypeptide. Derivatives may further include molecules formed by the deletion of one or more chemical groups which are naturally attached to the huE3 α polypeptide.

For example, the polypeptides may be modified by the covalent attachment of one or more polymers, including, but not limited to, water soluble polymers, N-linked or O-linked carbohydrates, sugars, phosphates, and/or other such molecules. For example, the polymer selected is typically water soluble so that the protein to which it is attached does not precipitate in an aqueous environment, such as a physiological environment. The polymer may be of any molecular weight, and may be branched or unbranched. Included within the scope of suitable polymers is a mixture of polymers. Preferably, for therapeutic use of the end-product preparation, the polymer will be pharmaceutically acceptable.

Suitable water soluble polymers or mixtures thereof include, but are not limited to, polyethylene glycol (PEG), monomethoxy-polyethylene glycol, dextran (such as low molecular weight dextran, of, for example about 6 kD), cellulose, or other carbohydrate based polymers, poly-(N-vinyl pyrrolidone) polyethylene glycol, propylene glycol homopolymers, a polypropylene oxide/ethylene oxide co-polymer, polyoxyethylated polyols (*e.g.*, glycerol) and polyvinyl alcohol. Also encompassed by the present invention are bifunctional PEG crosslinking molecules which may be used to prepare covalently attached huE3 α multimers.

For the acylation reactions, the polymer(s) selected should have a single reactive ester group. For reductive alkylation, the polymer(s) selected should have a single reactive aldehyde group. A reactive aldehyde is, for example, polyethylene glycol propionaldehyde, which is water stable, or mono C₁-C₁₀ alkoxy or aryloxy derivatives thereof (*see* U.S. Patent No. 5,252,714).

The pegylation of huE3 α polypeptides may be carried out by any of the pegylation reactions known in the art, as described for example in the following references: Francis *et al.*, *Focus on Growth Factors*, 3: 4-10, 1992; EP 0154316; EP 0401384 and U.S. Patent No. 4,179,337. Pegylation may be carried out via an acylation

reaction or an alkylation reaction with a reactive polyethylene glycol molecule (or an analogous reactive water-soluble polymer) as described herein.

Polyethylene glycol (PEG) is a water-soluble polymer suitable for use herein. As used herein, the terms "polyethylene glycol" and "PEG" are meant to encompass any of the forms of PEG that have been used to derivatize proteins, including mono-(C₁-C₁₀) alkoxy- or aryloxy-polyethylene glycol.

In general, chemical derivatization may be performed under any suitable conditions used to react a biologically active substance with an activated polymer molecule. Methods for preparing pegylated huE3 α polypeptides will generally comprise the steps of (a) reacting the polypeptide with polyethylene glycol (such as a reactive ester or aldehyde derivative of PEG) under conditions whereby huE3 α polypeptide becomes attached to one or more PEG groups, and (b) obtaining the reaction product(s). In general, the optimal reaction conditions for the acylation reactions will be determined based on known parameters and the desired result. For example, the larger the ratio of PEG:protein, the greater the percentage of poly-pegylated product. In one embodiment, the huE3 α polypeptide derivative may have a single PEG moiety at the amino terminus. See, for example, U.S. Patent No. 5,234,784.

Generally, conditions which may be alleviated or modulated by the administration of the present huE3 α polypeptide derivative include those described herein for huE3 α polypeptides. However, the huE3 α polypeptide derivative disclosed herein may have additional activities, enhanced or reduced biological activity, or other characteristics, such as increased or decreased half-life, as compared to the non-derivatized molecules.

The terms "biologically active huE3 α polypeptides", "biologically active huE3 α polypeptide fragments", "biologically active huE3 α polypeptide variants", and "biologically active huE3 α polypeptide derivatives" refer to huE3 α I or huE3 α II polypeptides having at least one activity characteristic of a human E3 α ubiquitin ligase, such as the activity of the polynucleotide set forth in SEQ ID NOS: 2 or 4. In general, huE3 α polypeptides, fragments, variants, and derivatives thereof, will have at least one activity characteristic of a huE3 α polypeptide such as depicted in SEQ ID NOS: 2 or 4. In addition, a huE3 α polypeptide may be active as an immunogen, that is, the polypeptide contains at least one epitope to which antibodies may be raised.

“Naturally occurring” or “native” when used in connection with biological materials such as nucleic acid molecules, polypeptides, host cells, and the like, refers to materials which are found in nature and are not manipulated by man. Similarly, “non-naturally occurring” or “non-native” as used herein refers to a material that is not found in nature or that has been structurally modified or synthesized by man.

The term “isolated polypeptide” refers to a polypeptide of the present invention that is free from at least one contaminating polypeptide that is found in its natural environment. Preferably, the isolated polypeptide is substantially free from any other contaminating mammalian polypeptides which would interfere with its therapeutic, preventative, or diagnostic use.

The term “ortholog” refers to a polypeptide that corresponds to a polypeptide identified from a different species that corresponds to huE3 α polypeptide amino acid sequence. For example, mouse and human E3 α polypeptides are considered orthologs.

The term “mature huE3 α polypeptide” refers to a polypeptide lacking a leader sequence. A mature polypeptide may also include other modifications such as proteolytic processing of the amino terminus (with or without a leader sequence) and/or the carboxy terminus, cleavage of a smaller polypeptide from a larger precursor, N-linked and/or O-linked glycosylation, and the like. An exemplary mature huE3 α polypeptide is depicted by SEQ ID NOS: 2 or 4.

The terms “effective amount” and “therapeutically effective amount” refer to the amount of a huE3 α polypeptide or huE3 α nucleic acid molecule used to support an observable level of one or more biological activities of the huE3 α polypeptides as set forth herein.

The term “selective binding agent” refers to a molecule or molecules having specificity for huE3 α molecules. Selective binding agents include antibodies, such as polyclonal antibodies, monoclonal antibodies (mAbs), chimeric antibodies, CDR-grafted antibodies, anti-idiotypic (anti-Id) antibodies to antibodies that can be labeled in soluble or bound form, as well as fragments, regions, or derivatives thereof

which are provided by known techniques, including, but not limited to enzymatic cleavage, peptide synthesis, or recombinant techniques.

As used herein, the terms, "specific" and "specificity" refer to the ability of the selective binding agents to bind to human huE3 α polypeptides. It will be appreciated, however, that the selective binding agents may also bind orthologs of huE3 α , polypeptides, that is, interspecies versions of E3 α , such as mouse and rat E3 α polypeptides. A preferred embodiment relates to antibodies that are highly specific to huE3 α polypeptides yet do not cross-react (that is, they fail to bind) with specificity to non-huE3 α polypeptides.

The term "antigen" refers to a molecule or a portion of a molecule capable of being bound by a selective binding agent, such as an antibody, which is additionally capable of inducing an animal to produce antibodies capable of binding to an epitope of that antigen. An antigen can have one or more epitopes. The specific binding reaction referred to above is meant to indicate that the antigen will react, in a highly selective manner, with its corresponding antibody and not with the multitude of other antibodies which can be evoked by other antigens.

Human E3 α polypeptides, fragments, variants, and derivatives may be used to prepare huE3 α selective binding agents using methods known in the art. Thus, antibodies and antibody fragments that bind huE3 α polypeptides are within the scope of the present invention. Antibody fragments include those portions of the antibody which bind to an epitope on the huE3 α polypeptide. Examples of such fragments include Fab and F(ab') fragments generated by enzymatic cleavage of full-length antibodies. Other binding fragments include those generated by recombinant DNA techniques, such as the expression of recombinant plasmids containing nucleic acid sequences encoding antibody variable regions. These antibodies may be, for example, polyclonal monospecific polyclonal, monoclonal, recombinant, chimeric, humanized, human, single chain, and/or bispecific.

Relatedness of Nucleic Acid Molecules and/or Polypeptides

The term "identity", as known in the art, refers to a relationship between the sequences of two or more polypeptide molecules or two or more nucleic

acid molecules, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between nucleic acid molecule or polypeptide sequences, as the case may be, as determined by the match between strings of two or more nucleotide or two or more amino acid sequences. "Identity" measures the percent of identical matches between two or more sequences with gap alignments (if any) addressed by a particular mathematical model or computer programs (*i.e.*, "algorithms").

The term "similarity" is a related concept, but in contrast to "identity", refers to a measure of similarity which includes both identical matches and conservative substitution matches. If two polypeptide sequences have, for example, 10/20 identical amino acids, and the remainder are all non-conservative substitutions, then the percent identity and similarity would both be 50%. If in the same example, there are 5 more positions where there are conservative substitutions, then the percent identity remains 50%, but the percent similarity would be 75% (15/20). Therefore, in cases where there are conservative substitutions, the degree of similarity between two polypeptide sequences will be higher than the percent identity between those two sequences.

The term "isolated nucleic acid molecule" refers to a nucleic acid molecule of the invention that (1) has been separated from at least about 50 percent of proteins, lipids, carbohydrates or other materials with which it is naturally found when total DNA is isolated from the source cells, (2) is not linked to all or a portion of a polynucleotide to which the "isolated nucleic acid molecule" is linked in nature, (3) is operably linked to a polynucleotide which it is not linked to in nature, or (4) does not occur in nature as part of a larger polynucleotide sequence. Preferably, the isolated nucleic acid molecule of the present invention is substantially free from any other contaminating nucleic acid molecule(s) or other contaminants that are found in its natural environment that would interfere with its use in polypeptide production or its therapeutic, diagnostic, prophylactic or research use.

The term "isolated polypeptide" refers to a polypeptide of the present invention that (1) has been separated from at least about 50 percent of polynucleotides, lipids, carbohydrates or other materials with which it is naturally found when isolated from the source cell, (2) is not linked (by covalent or noncovalent interaction) to all or

a portion of a polypeptide to which the "isolated polypeptide" is linked in nature, (3) is operably linked (by covalent or noncovalent interaction) to a polypeptide with which it is not linked in nature, or (4) does not occur in nature. Preferably, the isolated polypeptide is substantially free from any other contaminating polypeptides or other contaminants that are found in its natural environment that would interfere with its therapeutic, diagnostic, prophylactic or research use.

The term "conservative amino acid substitution" refers to a substitution of a native amino acid residue with a nonnative residue such that there is little or no effect on the polarity or charge of the amino acid residue at that position. For example, a conservative substitution results from the replacement of a non-polar residue in a polypeptide with any other non-polar residue. Furthermore, any native residue in the polypeptide may also be substituted with alanine, as has been previously described for "alanine scanning mutagenesis." General rules for conservative amino acid substitutions are set forth in Table II.

Table II
Amino Acid Substitutions

	Original Residues	Exemplary Substitutions	Preferred Substitutions
5	Ala	Val, Leu, Ile	Val
	Arg	Lys, Gln, Asn	Lys
	Asn	Gln	Gln
	Asp	Glu	Glu
10	Cys	Ser, Ala	Ser
	Gln	Asn	Asn
	Glu	Asp	Asp
	Gly	Pro, Ala	Ala
	His	Asn, Gln, Lys, Arg	Arg
15	Ile	Leu, Val, Met, Ala, Phe, Norleucine	Leu
	Leu	Norleucine, Ile, Val, Met, Ala, Phe	Ile
	Lys	Arg, 1,4 Diamino-butyric Acid, Gln, Asn	Arg
	Met	Leu, Phe, Ile	Leu
	Phe	Leu, Val, Ile, Ala, Tyr	Leu
20	Pro	Ala	Gly
	Ser	Thr, Ala, Cys	Thr
	Thr	Ser	Ser
	Trp	Tyr, Phe	Tyr
	Tyr	Trp, Phe, Thr, Ser	Phe
25	Val	Ile, Met, Leu, Phe, Ala, Norleucine	Leu

Conservative amino acid substitutions also encompass non-naturally occurring amino acid residues which are typically incorporated by chemical peptide synthesis rather than by synthesis in biological systems. These include peptidomimetics, and other reversed or inverted forms of amino acid moieties. It will be appreciated by those skilled in the art the nucleic acid and polypeptide molecules

described herein may be chemically synthesized as well as produced by recombinant means.

Conservative modifications to the amino acid sequence (and the corresponding modifications to the encoding nucleotides) will produce huE3 α polypeptides having functional and chemical characteristics similar to those of naturally occurring huE3 α polypeptides. In contrast, substantial modifications in the functional and/or chemical characteristics of huE3 α polypeptides may be accomplished by selecting substitutions that differ significantly in their effect on maintaining (a) the structure of the molecular backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain. Naturally occurring residues may be divided into classes based on common side chain properties:

- 1) hydrophobic: norleucine, Met, Ala, Val, Leu, Ile;
- 2) neutral hydrophilic: Cys, Ser, Thr, Asn, Gln;
- 3) acidic: Asp, Glu;
- 4) basic: His, Lys, Arg;
- 5) residues that influence chain orientation: Gly, Pro; and
- 6) aromatic: Trp, Tyr, Phe.

Non-conservative substitutions may involve the exchange of a member of one of these classes for a member from another class. Such substituted residues may be introduced into regions of the human E3 α polypeptide that are homologous with non-human E3 α polypeptides, or into the non-homologous regions of the molecule.

Identity and similarity of related nucleic acid molecules and polypeptides can be readily calculated by known methods. Such methods include, but are not limited to, those described in *Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data*, Part 1, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987; *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M.

Stockton Press, New York, 1991; and Carillo *et al.*, *SIAM J. Applied Math.*, 48: 1073, 1988.

Preferred methods to determine identity and/or similarity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are described in publicly available computer programs. Preferred computer program methods to determine identity and similarity between two sequences include, but are not limited to, the GCG program package, including GAP (Devereux *et al.*, *Nucl. Acid. Res.*, 12: 387, 1984; Genetics Computer Group, University of Wisconsin, Madison, WI), BLASTP, BLASTN, and FASTA (Altschul *et al.*, *J. Mol. Biol.*, 215: 403-410, 1990). The BLASTX program is publicly available from the National Center for Biotechnology Information (NCBI) and other sources (*BLAST Manual*, Altschul *et al.* NCB/NLM/NIH Bethesda, MD 20894; Altschul *et al.*, *supra*). The well known Smith Waterman algorithm may also be used to determine identity.

Certain alignment schemes for aligning two amino acid sequences may result in the matching of only a short region of the two sequences, and this small aligned region may have very high sequence identity even though there is no significant relationship between the two full length sequences. Accordingly, in a preferred embodiment, the selected alignment method (GAP program) will result in an alignment that spans at least 50 contiguous amino acids of the target polypeptide.

For example, using the computer algorithm GAP (Genetics Computer Group, University of Wisconsin, Madison, WI), two polypeptides for which the percent sequence identity is to be determined are aligned for optimal matching of their respective amino acids (the "matched span", as determined by the algorithm). A gap opening penalty (which is calculated as 3x the average diagonal; the "average diagonal" is the average of the diagonal of the comparison matrix being used; the "diagonal" is the score or number assigned to each perfect amino acid match by the particular comparison matrix) and a gap extension penalty (which is usually 1/10 times the gap opening penalty), as well as a comparison matrix such as PAM 250 or BLOSUM 62 are used in conjunction with the algorithm. A standard comparison matrix (see Dayhoff *et al.*, *Atlas of Protein Sequence and Structure*, vol. 5, supp.3 (1978) for the PAM 250 comparison matrix; Henikoff *et al.*, *Proc. Natl. Acad. Sci USA*, 89: 10915-10919, 1992 for the BLOSUM 62 comparison matrix) is also used by the algorithm.

Preferred parameters for a polypeptide sequence comparison include the following:

Algorithm: Needleman *et al.*, *J. Mol. Biol.*, 48, 443-453, 1970;

Comparison matrix: BLOSUM 62 from Henikoff *et al.*, *Proc. Natl. Acad. Sci.*
USA, 89: 10915-10919, 1992)

Gap Penalty: 12

Gap Length Penalty: 4

Threshold of Similarity: 0

The GAP program is useful with the above parameters. The
aforementioned parameters are the default parameters for polypeptide comparisons
(along with no penalty for end gaps) using the GAP algorithm.

Preferred parameters for nucleic acid molecule sequence comparisons
include the following:

Algorithm: Needleman *et al.*, *J. Mol Biol.*, 48: 443-453, 1970;

Comparison matrix: matches = +10, mismatch = 0

Gap Penalty: 50

Gap Length Penalty: 3

The GAP program is also useful with the above parameters. The
aforementioned parameters are the default parameters for nucleic acid molecule
comparisons.

Other exemplary algorithms, gap opening penalties, gap extension
penalties, comparison matrices, thresholds of similarity, etc. may be used by those of
skill in the art, including those set forth in the Program Manual, Wisconsin Package,
Version 9, September, 1997. The particular choices to be made will be apparent to
those of skill in the art and will depend on the specific comparison to be made, such
as DNA to DNA, protein to protein, protein to DNA; and additionally, whether the
comparison is between given pairs of sequences (in which case GAP or BestFit are
generally preferred) or between one sequence and a large database of sequences (in
which case FASTA or BLASTA are preferred).

Synthesis

It will be appreciated by those skilled in the art the nucleic acid and polypeptide molecules described herein may be produced by recombinant and other means.

Nucleic Acid Molecules

Recombinant DNA methods used herein are generally those set forth in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989), and/or Ausubel *et al.*, eds., *Current Protocols in Molecular Biology*, Green Publishers Inc. and Wiley and Sons, NY (1994). The present invention provides for nucleic acid molecules as described herein and methods for obtaining the molecules. Human E3 α (huE3 α) refers to the nucleotide sequence of either huE3 α I or huE3 α II. A gene or cDNA encoding a huE3 α polypeptide or fragment thereof may be obtained by hybridization screening of a genomic or cDNA library, or by PCR amplification. Where a gene encoding a huE3 α polypeptide has been identified from one species, all or a portion of that gene may be used as a probe to identify corresponding genes from other species (orthologs) or related genes from the same species (homologs). The probes or primers may be used to screen cDNA libraries from various tissue sources believed to express the huE3 α gene. In addition, part or all of a nucleic acid molecule having the sequence as set forth in SEQ ID NOS: 1 or 3 may be used to screen a genomic library to identify and isolate a gene encoding a huE3 α polypeptide. Typically, conditions of moderate or high stringency will be employed for screening to minimize the number of false positives obtained from the screen.

Nucleic acid molecules encoding huE3 α polypeptides may also be identified by expression cloning which employs the detection of positive clones based upon a property of the expressed protein. Typically, nucleic acid libraries are screened by the binding of an antibody or other binding partner (*e.g.*, receptor or ligand) to cloned proteins which are expressed and displayed on a host cell surface. The antibody or binding partner is modified with a detectable label to identify those cells expressing the desired clone.

Additional methods of predicting secondary structure include “threading” (Jones *et al.*, *Current Opin. Struct. Biol.*, 7(3):377-87 (1997); Sippl *et al.*, *Structure*, 4(1):15-9 (1996)), “profile analysis” (Bowie *et al.*, *Science*, 253:164-170 (1991); Gribskov *et al.*, *Meth. Enzym.*, 183:146-159 (1990); Gribskov *et al.*, *Proc. Nat. Acad. Sci.*, 84(13):4355-4358 (1987)), and “evolutionary linkage” (See Home, *supra*, and Brenner, *supra* 1997).

Another means of preparing a nucleic acid molecule encoding a huE3 α polypeptide, including a fragment or variant, is chemical synthesis using methods well known to the skilled artisan such as those described by Engels *et al.*, *Angew. Chem. Intl. Ed.*, 28: 716-734, 1989. These methods include, *inter alia*, the phosphotriester, phosphoramidite, and H-phosphonate methods for nucleic acid synthesis. A preferred method for such chemical synthesis is polymer-supported synthesis using standard phosphoramidite chemistry. Typically, the DNA encoding the huE3 α polypeptide will be several hundred nucleotides in length. Nucleic acids larger than about 100 nucleotides can be synthesized as several fragments using these methods. The fragments can then be ligated together to form the full length huE3 α polypeptide. Usually, the DNA fragment encoding the amino terminus of the polypeptide will have an ATG, which encodes a methionine residue. This methionine may or may not be present on the mature form of the huE3 α polypeptide, depending on whether the polypeptide produced in the host cell is designed to be secreted from that cell.

In some cases, it may be desirable to prepare nucleic acid molecules encoding huE3 α polypeptide variants. Nucleic acid molecules encoding variants may be produced using site directed mutagenesis, PCR amplification, or other appropriate methods, where the primer(s) have the desired point mutations (see Sambrook *et al.*, *supra*, and Ausubel *et al.*, *supra*, for descriptions of mutagenesis techniques). Chemical synthesis using methods described by Engels *et al.*, *supra*, may also be used to prepare such variants. Other methods known to the skilled artisan may be used as well.

In certain embodiments, nucleic acid variants contain codons which have been altered for the optimal expression of a huE3 α polypeptide in a given host cell. Particular codon alterations will depend upon the huE3 α polypeptide(s) and host cell(s) selected for expression. Such “codon optimization” can be carried out by a variety of methods, for example, by selecting codons which are preferred for use in highly

expressed genes in a given host cell. Computer algorithms which incorporate codon frequency tables such as "Ecohigh.cod" for codon preference of highly expressed bacterial genes may be used and are provided by the University of Wisconsin Package Version 9.0, Genetics Computer Group, Madison, WI. Other useful codon frequency tables include "Celegans_high.cod", "Celegans_low.cod", "Drosophila_high.cod", "Human_high.cod", "Maize_high.cod", and "Yeast_high.cod".

In other embodiments, nucleic acid molecules encode huE3 α variants with conservative amino acid substitutions as described herein, huE3 α variants comprising an addition and/or a deletion of one or more N-linked or O-linked glycosylation sites, huE3 α variants having deletions and/or substitutions of one or more cysteine residues, or huE3 α polypeptide fragments as described herein. In addition, nucleic acid molecules may encode any combination of huE3 α variants, fragments, and fusion polypeptides described herein.

Vectors and Host Cells

A nucleic acid molecule encoding a huE3 α polypeptide is inserted into an appropriate expression vector using standard ligation techniques wherein huE3 α refers to either the polypeptide sequence of huE3 α I or huE3 α II. The vector is typically selected to be functional in the particular host cell employed (*i.e.*, the vector is compatible with the host cell machinery such that amplification of the gene and/or expression of the gene can occur). A nucleic acid molecule encoding a huE3 α polypeptide may be amplified/expressed in prokaryotic, yeast, insect (baculovirus systems), and/or eukaryotic host cells. Selection of the host cell will depend in part on whether a huE3 α polypeptide is to be post-translationally modified (*e.g.*, glycosylated and/or phosphorylated). If so, yeast, insect, or mammalian host cells are preferable. For a review of expression vectors, see *Meth. Enz.*, v.185, D.V. Goeddel, ed. Academic Press Inc., San Diego, CA (1990).

Typically, expression vectors used in any of the host cells will contain sequences for plasmid maintenance and for cloning and expression of exogenous nucleotide sequences. Such sequences, collectively referred to as "flanking sequences" in certain embodiments will typically include one or more of the following nucleotides: a promoter, one or more enhancer sequences, an origin of replication, a transcriptional termination sequence, a complete intron sequence containing a donor and acceptor

splice site, a sequence encoding a leader sequence for polypeptide secretion, a ribosome binding site, a polyadenylation sequence, a polylinker region for inserting the nucleic acid encoding the polypeptide to be expressed, and a selectable marker element. Each of these sequences is discussed below.

5 Optionally, the vector may contain a "tag"-encoding sequence, *i.e.*, an oligonucleotide sequence located at the 5' or 3' end of the huE3 α polypeptide coding sequence; the oligonucleotide molecule encodes polyHis (such as hexaHis), or another "tag" such as FLAG, HA (hemagglutinin influenza virus) or *myc* for which commercially available antibodies exist. This tag is typically fused to the polypeptide upon expression of the polypeptide, and can serve as a means for affinity purification of the huE3 α polypeptide from the host cell. Affinity purification can be accomplished, for example, by column chromatography using antibodies against the tag as an affinity matrix. Optionally, the tag can subsequently be removed from the purified huE3 α polypeptide by various means such as using certain peptidases for cleavage.

15 Flanking sequences may be homologous (*i.e.*, from the same species and/or strain as the host cell), heterologous (*i.e.*, from a species other than the host cell species or strain), hybrid (*i.e.*, a combination of flanking sequences from more than one source), or synthetic, or the flanking sequences may be native sequences which normally function to regulate huE3 α polypeptide expression. As such, the source of a flanking sequence may be any prokaryotic or eukaryotic organism, any vertebrate or invertebrate organism, or any plant, provided that the flanking sequences is functional in, and can be activated by, the host cell machinery.

25 The flanking sequences useful in the vectors of this invention may be obtained by any of several methods well known in the art. Typically, flanking sequences useful herein other than endogenous huE3 α gene flanking sequences will have been previously identified by mapping and/or by restriction endonuclease digestion and can thus be isolated from the proper tissue source using the appropriate restriction endonucleases. In some cases, the full nucleotide sequence of one or more flanking sequence may be known. Here, the flanking sequence may be synthesized using the methods described herein for nucleic acid synthesis or cloning.

Where all or only a portion of the flanking sequence is known, it may be obtained using PCR and/or by screening a genomic library with suitable oligonucleotide and/or flanking sequence fragments from the same or another species. Where the flanking sequence is not known, a fragment of DNA containing a flanking sequence may be isolated from a larger piece of DNA that may contain, for example, a coding sequence or even another gene or genes. Isolation may be accomplished by restriction endonuclease digestion to produce the proper DNA fragment followed by isolation using agarose gel purification, Qiagen® column chromatography (Chatsworth, CA), or other methods known to the skilled artisan. The selection of suitable enzymes to accomplish this purpose will be readily apparent to one of ordinary skill in the art.

An origin of replication is typically a part of those prokaryotic expression vectors purchased commercially, and the origin aids in the amplification of the vector in a host cell. Amplification of the vector to a certain copy number can, in some cases, be important for the optimal expression of the huE3 α polypeptide. If the vector of choice does not contain an origin of replication site, one may be chemically synthesized based on a known sequence, and ligated into the vector. For example, the origin of replication from the plasmid pBR322 (Product No. 303-3s, New England Biolabs, Beverly, MA) is suitable for most gram-negative bacteria and various origins (*e.g.*; SV40, polyoma, adenovirus, vesicular stomatitis virus (VSV) or papillomaviruses such as HPV or BPV) are useful for cloning vectors in mammalian cells. Generally, the origin of replication component is not needed for mammalian expression vectors (for example, the SV40 origin is often used only because it contains the early promoter).

A transcription termination sequence is typically located 3' of the end of a polypeptide coding region and serves to terminate transcription. Usually, a transcription termination sequence in prokaryotic cells is a G-C rich fragment followed by a poly T sequence. While the sequence is easily cloned from a library or even purchased commercially as part of a vector, it can also be readily synthesized using methods for nucleic acid synthesis such as those described herein.

A selectable marker gene element encodes a protein necessary for the survival and growth of a host cell grown in a selective culture medium. Typical selection marker genes encode proteins that (a) confer resistance to antibiotics or other

toxins, *e.g.*, ampicillin, tetracycline, or kanamycin for prokaryotic host cells, (b) complement auxotrophic deficiencies of the cell; or (c) supply critical nutrients not available from complex media. Preferred selectable markers are the kanamycin resistance gene, the ampicillin resistance gene, and the tetracycline resistance gene.

5 A neomycin resistance gene may also be used for selection in prokaryotic and eukaryotic host cells.

Other selection genes may be used to amplify the gene which will be expressed. Amplification is the process wherein genes which are in greater demand for the production of a protein critical for growth are reiterated in tandem within the

10 chromosomes of successive generations of recombinant cells. Examples of suitable selectable markers for mammalian cells include dihydrofolate reductase (DHFR) and thymidine kinase. The mammalian cell transformants are placed under selection pressure which only the transformants are uniquely adapted to survive by virtue of the selection gene present in the vector. Selection pressure is imposed by culturing the

15 transformed cells under conditions in which the concentration of selection agent in the medium is successively changed, thereby leading to the amplification of both the selection gene and the DNA that encodes huE3 α polypeptides. As a result, increased quantities of huE3 α polypeptides are synthesized from the amplified DNA.

A ribosome binding site is usually necessary for translation initiation of mRNA and is characterized by a Shine-Dalgarno sequence (prokaryotes) or a Kozak

20 sequence (eukaryotes). The element is typically located 3' to the promoter and 5' to the coding sequence of the huE3 α polypeptide to be expressed. The Shine-Dalgarno sequence is varied but is typically a polypurine (*i.e.*, having a high A-G content). Many Shine-Dalgarno sequences have been identified, each of which can be readily

25 synthesized using methods set forth above and used in a prokaryotic vector.

A leader, or signal, sequence may be used to direct a huE3 α polypeptide out of the host cell. Typically, a nucleotide sequence encoding the signal sequence is positioned in the coding region of the huE3 α nucleic acid molecule, or directly at the

30 5' end of the huE3 α polypeptide coding region. Many signal sequences have been identified, and any of those that are functional in the selected host cell may be used in conjunction with the huE3 α nucleic acid molecule. Therefore, a signal sequence may be homologous (naturally occurring) or heterologous to the huE3 α gene or cDNA.

Additionally, a signal sequence may be chemically synthesized using methods described herein. In most cases, the secretion of a huE3 α polypeptide from the host cell via the presence of a signal peptide will result in the removal of the signal peptide from the huE3 α polypeptide. The signal sequence may be a component of the vector, or it may be a part of huE3 α DNA that is inserted into the vector.

Included within the scope of this invention is the use of either a nucleotide sequence encoding a native huE3 α signal sequence joined to a huE3 α polypeptide coding region or a nucleotide sequence encoding a heterologous signal sequence joined to a huE3 α polypeptide coding region. The heterologous signal sequence selected should be one that is recognized and processed, *i.e.*, cleaved by a signal peptidase, by the host cell. For prokaryotic host cells that do not recognize and process the native huE3 α signal sequence, the signal sequence is substituted by a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, penicillinase, or heat-stable enterotoxin II leaders. For yeast secretion, the native huE3 α signal sequence may be substituted by the yeast invertase, alpha factor, or acid phosphatase leaders. In mammalian cell expression the native signal sequence is satisfactory, although other mammalian signal sequences may be suitable.

In some cases, such as where glycosylation is desired in a eukaryotic host cell expression system, one may manipulate the various presequences to improve glycosylation or yield. For example, one may alter the peptidase cleavage site of a particular signal peptide, or add presequences, which also may affect glycosylation. The final protein product may have, in the -1 position (relative to the first amino acid of the mature protein) one or more additional amino acids incident to expression, which may not have been totally removed. For example, the final protein product may have one or two amino acid residues found in the peptidase cleavage site, attached to the N-terminus. Alternatively, use of some enzyme cleavage sites may result in a slightly truncated form of the desired huE3 α polypeptide, if the enzyme cuts at such area within the mature polypeptide.

In many cases, transcription of a nucleic acid molecule is increased by the presence of one or more introns in the vector; this is particularly true where a polypeptide is produced in eukaryotic host cells, especially mammalian host cells. The introns used may be naturally occurring within the huE3 α gene, especially where the

gene used is a full length genomic sequence or a fragment thereof. Where the intron is not naturally occurring within the gene (as for most cDNAs), the intron(s) may be obtained from another source. The position of the intron with respect to flanking sequences and the huE3 α gene is generally important, as the intron must be expressed to be effective. Thus, when a huE3 α cDNA molecule is being expressed, the preferred position for the intron is 3' to the transcription start site, and 5' to the polyA transcription termination sequence. Preferably, the intron or introns will be located on one side or the other (*i.e.*, 5' or 3') of the cDNA such that it does not interrupt the coding sequence. Any intron from any source, including any viral, prokaryotic and eukaryotic (plant or animal) organisms, may be used to practice this invention, provided that it is compatible with the host cell(s) into which it is inserted. Also included herein are synthetic introns. Optionally, more than one intron may be used in the vector.

The expression and cloning vectors of the present invention will each typically contain a promoter that is recognized by the host organism and operably linked to the molecule encoding a huE3 α polypeptide. Promoters are untranscribed sequences located upstream (5') to the start codon of a structural gene (generally within about 100 to 1000 bp) that control the transcription and translation of the structural gene. Promoters are conventionally grouped into one of two classes, inducible promoters and constitutive promoters. Inducible promoters initiate increased levels of transcription from DNA under their control in response to some change in culture conditions, such as the presence or absence of a nutrient or a change in temperature. Constitutive promoters, on the other hand, initiate continual gene product production; that is, there is little or no control over gene expression. A large number of promoters, recognized by a variety of potential host cells, are well known. A suitable promoter is operably linked to the DNA encoding a huE3 α polypeptide by removing the promoter from the source DNA by restriction enzyme digestion and inserting the desired promoter sequence into the vector. The native huE3 α promoter sequence may be used to direct amplification and/or expression of huE3 α DNA. A heterologous promoter is preferred, however, if it permits greater transcription and higher yields of the expressed protein as compared to the native promoter, and if it is compatible with the host cell system that has been selected for use.

Promoters suitable for use with prokaryotic hosts include the beta-lactamase and lactose promoter systems; alkaline phosphatase, a tryptophan (trp) promoter system; and hybrid promoters such as the tac promoter. Other known bacterial promoters are also suitable. Their sequences have been published, thereby enabling one skilled in the art to ligate them to the desired DNA sequence(s), using linkers or adapters as needed to supply any useful restriction sites.

Suitable promoters for use with yeast hosts are also well known in the art. Yeast enhancers are advantageously used with yeast promoters. Suitable promoters for use with mammalian host cells are well known and include, but are not limited to, those obtained from the genomes of viruses such as polyoma virus, fowl pox virus, adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus (CMV), a retrovirus, hepatitis-B virus and most preferably Simian Virus 40 (SV40). Other suitable mammalian promoters include heterologous mammalian promoters, *e.g.*, heat-shock promoters and the actin promoter.

Additional promoters which may be of interest in controlling huE3 α gene transcription include, but are not limited to: the SV40 early promoter region (Bernoist and Chambon, *Nature*, 290: 304-310, 1981); the CMV promoter; the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto *et al.*, *Cell*, 22: 787-797, 1980); the herpes thymidine kinase promoter (Wagner *et al.*, *Proc. Natl. Acad. Sci. USA*, 78: 144-1445, 1981); the regulatory sequences of the metallothionine gene (Brinster *et al.*, *Nature*, 296: 39-42, 1982); prokaryotic expression vectors such as the beta-lactamase promoter (Villa-Kamaroff, *et al.*, *Proc. Natl. Acad. Sci. USA*, 75: 3727-3731, 1978); or the tac promoter (DeBoer, *et al.*, *Proc. Natl. Acad. Sci. USA*, 80: 21-25, 1983). Also of interest are the following animal transcriptional control regions, which exhibit tissue specificity and have been utilized in transgenic animals: the elastase I gene control region which is active in pancreatic acinar cells (Swift *et al.*, *Cell*, 38: 639-646, 1984; Ornitz *et al.*, *Cold Spring Harbor Symp. Quant. Biol.*, 50: 399-409, 1986; MacDonald, *Hepatology*, 7: 425-515, 1987); the insulin gene control region which is active in pancreatic beta cells (Hanahan, *Nature*, 315: 115-122, 1985); the immunoglobulin gene control region which is active in lymphoid cells (Grosschedl *et al.*, *Cell*, 38: 647-658 (1984); Adames *et al.*, *Nature*, 318: 533-538 (1985); Alexander *et al.*, *Mol. Cell. Biol.*, 7: 1436-1444,

1987); the mouse mammary tumor virus control region which is active in testicular, breast, lymphoid and mast cells (Leder *et al.*, *Cell*, 45: 485-495, 1986); the albumin gene control region which is active in liver (Pinkert *et al.*, *Genes and Devel.*, 1: 268-276, 1987); the alphafetoprotein gene control region which is active in liver (Krumlauf *et al.*, *Mol. Cell. Biol.*, 5: 1639-1648, 1985; Hammer *et al.*, *Science*, 235: 53-58, 1987); the alpha 1-antitrypsin gene control region which is active in the liver (Kelsey *et al.*, *Genes and Devel.*, 1: 161-171, 1987); the beta-globin gene control region which is active in myeloid cells (Mogram *et al.*, *Nature*, 315: 338-340, 1985; Kollias *et al.*, *Cell*, 46: 89-94, 1986); the myelin basic protein gene control region which is active in oligodendrocyte cells in the brain (Readhead *et al.*, *Cell*, 48: 703-712, 1987); the myosin light chain-2 gene control region which is active in skeletal muscle (Sani, *Nature*, 314: 283-286, 1985); and the gonadotropic releasing hormone gene control region which is active in the hypothalamus (Mason *et al.*, *Science*, 234: 1372-1378, 1986).

An enhancer sequence may be inserted into the vector to increase the transcription of a DNA encoding a huE3 α polypeptide of the present invention by higher eukaryotes. Enhancers are cis-acting elements of DNA, usually about 10-300 bp in length, that act on the promoter to increase its transcription. Enhancers are relatively orientation and position independent. They have been found 5' and 3' to the transcription unit. Several enhancer sequences available from mammalian genes are known (*e.g.*, globin, elastase, albumin, alpha-feto-protein and insulin). Typically, however, an enhancer from a virus will be used. The SV40 enhancer, the cytomegalovirus early promoter enhancer, the polyoma enhancer, and adenovirus enhancers are exemplary enhancing elements for the activation of eukaryotic promoters. While an enhancer may be spliced into the vector at a position 5' or 3' to huE3 α DNA, it is typically located at a site 5' from the promoter.

Expression vectors of the invention may be constructed from a starting vector such as a commercially available vector. Such vectors may or may not contain all of the desired flanking sequences. Where one or more of the desired flanking sequences are not already present in the vector, they may be individually obtained and ligated into the vector. Methods used for obtaining each of the flanking sequences are well known to one skilled in the art.

Preferred vectors for practicing this invention are those which are compatible with bacterial, insect, and mammalian host cells. Such vectors include, *inter alia*, pCRII, pCR3, and pcDNA3.1 (Invitrogen Company, Carlsbad, CA), pBSII (Stratagene Company, La Jolla, CA), pET15 (Novagen, Madison, WI), pGEX (Pharmacia Biotech, Piscataway, NJ), pEGFP-N2 (Clontech, Palo Alto, CA), pETL (BlueBacII; Invitrogen), pDSR-alpha (PCT Publication No. WO90/14363) and pFastBacDual (Gibco/BRL, Grand Island, NY).

Additional suitable vectors include, but are not limited to, cosmids, plasmids, or modified viruses, but it will be appreciated that the vector system must be compatible with the selected host cell. Such vectors include, but are not limited to plasmids such as Bluescript[®] plasmid derivatives (a high copy number ColE1-based phagemid, Stratagene Cloning Systems Inc., La Jolla CA), PCR cloning plasmids designed for cloning Taq-amplified PCR products (e.g., TOPO[™] TA Cloning[®] Kit, PCR2.1[®] plasmid derivatives, Invitrogen, Carlsbad, CA), and mammalian, yeast, or virus vectors such as a baculovirus expression system (pBacPAK plasmid derivatives, Clontech, Palo Alto, CA). The recombinant molecules can be introduced into host cells via transformation, transfection, infection, electroporation, or other known techniques.

After the vector has been constructed and a nucleic acid molecule encoding a huE3 α polypeptide has been inserted into the proper site of the vector, the completed vector may be inserted into a suitable host cell for amplification and/or polypeptide expression. Host cells may be prokaryotic host cells (such as *E. coli*) or eukaryotic host cells (such as a yeast cell, an insect cell, or a vertebrate cell). The host cell, when cultured under appropriate conditions, synthesizes a huE3 α polypeptide which can subsequently be collected from the culture medium (if the host cell secretes it into the medium) or directly from the host cell producing it (if it is not secreted). The selection of an appropriate host cell will depend upon various factors, such as desired expression levels, polypeptide modifications that are desirable or necessary for activity, such as glycosylation or phosphorylation, and ease of folding into a biologically active molecule.

A number of suitable host cells are known in the art and many are available from the American Type Culture Collection (ATCC), 10801 Univeristy Boulavard, Manassas, VA 20110-2209. Examples include, but are not limited to, mammalian cells, such as Chinese hamster ovary cells (CHO) (ATCC No. CCL61) CHO DHFR-cells (Urlaub *et al.*, *Proc. Natl. Acad. Sci. USA*, 97: 4216-4220, 1980), human

embryonic kidney (HEK) 293 or 293T cells (ATCC No. CRL1573), or 3T3 cells (ATCC No. CCL92). The selection of suitable mammalian host cells and methods for transformation, culture, amplification, screening and product production and purification are known in the art. Other suitable mammalian cell lines, are the monkey COS-1 (ATCC No. CRL1650) and COS-7 cell lines (ATCC No. CRL1651), and the CV-1 cell line (ATCC No. CCL70). Further exemplary mammalian host cells include primate cell lines and rodent cell lines, including transformed cell lines. Normal diploid cells, cell strains derived from *in vitro* culture of primary tissue, as well as primary explants, are also suitable. Candidate cells may be genotypically deficient in the selection gene, or may contain a dominantly acting selection gene. Other suitable mammalian cell lines include but are not limited to, mouse neuroblastoma N2A cells, HeLa, mouse L-929 cells, 3T3 lines derived from Swiss, Balb-c or NIH mice, BHK or HaK hamster cell lines, which are available from the American Type Culture Collection, Manassas, VA. Each of these cell lines is known by and available to those skilled in the art of protein expression.

Similarly useful as host cells suitable for the present invention are bacterial cells. For example, the various strains of *E. coli* (*e.g.*, HB101, (ATCC No. 33694) DH5 α , DH10, and MC1061 (ATCC No. 53338)) are well-known as host cells in the field of biotechnology. Various strains of *B. subtilis*, *Pseudomonas spp.*, other *Bacillus spp.*, *Streptomyces spp.*, and the like may also be employed in this method.

Many strains of yeast cells known to those skilled in the art are also available as host cells for the expression of the polypeptides of the present invention. Preferred yeast cells include, for example, *Saccharomyces cerevisiae* and *Pichia pastoris*.

Additionally, where desired, insect cell systems may be utilized in the methods of the present invention. Such systems are described for example in Kitts *et al.*, *Biotechniques*, 14: 810-817 (1993); Lucklow, *Curr. Opin. Biotechnol.*, 4: 564-572, 1993; and Lucklow *et al. J. Virol.*, 67: 4566-4579, 1993. Preferred insect cells are Sf-9 and Hi5 (Invitrogen, Carlsbad, CA).

The transformation of an expression vector for a huE3 α polypeptide into a selected host cell may be accomplished by well known methods including methods such as transfection, infection, calcium chloride, electroporation, microinjection, lipofection or the DEAE-dextran method or other known techniques. The method selected will in part be a function of the type of host cell to be used. These methods and

other suitable methods are well known to the skilled artisan, and are set forth, for example, in Sambrook *et al.*, *supra*.

One may also use transgenic animals to express glycosylated huE3 α polypeptides. For example, one may use a transgenic milk-producing animal (a cow or goat, for example) and obtain the present glycosylated polypeptide in the animal milk. One may also use plants to produce huE3 α polypeptides, however, in general, the glycosylation occurring in plants is different from that produced in mammalian cells, and may result in a glycosylated product which is not suitable for human therapeutic use.

Polypeptide Production

Host cells comprising a huE3 α expression vector may be cultured using standard media well known to the skilled artisan. The huE3 α expression vector refers to a vector which expresses either huE3 α I or huE3 α II. The media will usually contain all nutrients necessary for the growth and survival of the cells. Suitable media for culturing *E. coli* cells include, for example, Luria Broth (LB) and/or Terrific Broth (TB). Suitable media for culturing eukaryotic cells are, Roswell Park Memorial Institute medium 1640 (RPMI 1640), Minimal Essential Medium (MEM), and/or Dulbecco's Modified Eagle Medium (DMEM), all of which may be supplemented with serum and/or growth factors as indicated by the particular cell line being cultured. A suitable medium for insect cultures is Grace's medium supplemented with yeastolate, lactalbumin hydrolysate, and/or fetal calf serum as necessary.

Typically, an antibiotic or other compound useful for selective growth of transformed cells is added as a supplement to the media. The compound to be used will be dictated by the selectable marker element present on the plasmid with which the host cell was transformed. For example, where the selectable marker element is kanamycin resistance, the compound added to the culture medium will be kanamycin. Other compounds for selective growth include ampicillin, tetracycline, and neomycin.

The amount of a huE3 α polypeptide produced by a host cell can be evaluated using standard methods known in the art. Such methods include, without limitation, Western blot analysis, SDS-polyacrylamide gel electrophoresis, non-denaturing gel electrophoresis, High Performance Liquid Chromatography (HPLC) separation, immunoprecipitation, and/or activity assays such as DNA binding gel shift assays.

If a huE3 α polypeptide has been designed to be secreted from the host cells, the majority of polypeptide may be found in the cell culture medium. If however, the huE3 α polypeptide is not secreted from the host cells, it will be present in the cytoplasm and/or the nucleus (for eukaryotic host cells) or in the cytosol (for bacterial host cells).

For a huE3 α polypeptide situated in the host cell cytoplasm and/or nucleus, the host cells are typically first disrupted mechanically or with a detergent to release the intracellular contents into a buffered solution. Human E3 α polypeptide can then be isolated from this solution.

The purification of a huE3 α polypeptide from solution can be accomplished using a variety of techniques. If the polypeptide has been synthesized such that it contains a tag such as Hexahistidine (huE3 α polypeptide/hexaHis) or other small peptide such as FLAG (Eastman Kodak Co., New Haven, CT) or *myc* (Invitrogen, Carlsbad, CA) at either its carboxyl or amino terminus, it may essentially be purified in a one-step process by passing the solution through an affinity column where the column matrix has a high affinity for the tag or for the polypeptide directly (*i.e.*, a monoclonal antibody specifically recognizing and binding to the huE3 α polypeptide). For example, polyhistidine binds with great affinity and specificity to nickel, thus an affinity column of nickel (such as the Qiagen[®] nickel columns) can be used for purification of huE3 α polypeptide/polyHis. See for example, Ausubel *et al.*, eds., *Current Protocols in Molecular Biology*, Section 10.11.8, John Wiley & Sons, New York (1993).

Where a huE3 α polypeptide is prepared without a tag attached, and no antibodies are available, other well known procedures for purification can be used. Such procedures include, without limitation, ion exchange chromatography, molecular sieve chromatography, High Performance Liquid Chromatography (HPLC), native gel electrophoresis in combination with gel elution, and preparative isoelectric focusing ("Isoprime" machine/technique, Hoefer Scientific, San Francisco, CA). In some cases, two or more of these techniques may be combined to achieve increased purity.

If a huE3 α polypeptide is produced intracellularly, the intracellular material (including inclusion bodies for gram-negative bacteria) can be extracted from the host cell using any standard technique known to the skilled artisan. For example, the host cells can be lysed to release the contents of the periplasm/cytoplasm by French press, homogenization, and/or sonication followed by centrifugation.

If a huE3 α polypeptide has formed inclusion bodies in the cytosol, the inclusion bodies can often bind to the inner and/or outer cellular membranes and thus will be found primarily in the pellet material after centrifugation. The pellet material can then be treated at pH extremes or with chaotropic agent such as a detergent, guanidine, guanidine derivatives, urea, or urea derivatives in the presence of a reducing agent such as dithiothreitol at alkaline pH or tris carboxyethyl phosphine at acid pH to release, break apart, and solubilize the inclusion bodies. The solubilized huE3 α polypeptide can then be analyzed using gel electrophoresis, immunoprecipitation or the like. If it is desired to isolate the huE3 α polypeptide, isolation may be accomplished using standard methods such as those described herein and in Marston *et al.*, *Meth. Enz.*, 182: 264-275 1990.

In some cases, a huE3 α polypeptide may not be biologically active upon isolation. Various methods for "refolding" or converting the polypeptide to its tertiary structure and generating disulfide linkages, can be used to restore biological activity. Such methods include exposing the solubilized polypeptide to a pH usually above 7 and in the presence of a particular concentration of a chaotrope. The selection of chaotrope is very similar to the choices used for inclusion body solubilization, but usually the chaotrope is used at a lower concentration and is not necessarily the same as chaotropes used for the solubilization. In most cases the refolding/oxidation solution will also contain a reducing agent or the reducing agent plus its oxidized form in a specific ratio to generate a particular redox potential allowing for disulfide shuffling to occur in the formation of the protein's cysteine bridge(s). Some of the commonly used redox couples include cysteine/cystamine, glutathione (GSH)/dithiobis GSH, cupric chloride, dithiothreitol (DTT)/ dithiane DTT, and 2-2mercaptoethanol(β ME)/dithi(β ME). A cosolvent may be used to increase the efficiency of the refolding, and the more common reagents used for this purpose include glycerol, polyethylene glycol of various molecular weights, arginine and the like.

If inclusion bodies are not formed to a significant degree upon expression of a huE3 α polypeptide, then the polypeptide will be found primarily in the supernatant after centrifugation of the cell homogenate. The polypeptide may be further isolated from the supernatant using methods such as those described herein.

In situations where it is preferable to partially or completely purify a huE3 α polypeptide such that it is partially or substantially free of contaminants, standard methods known to those skilled in the art may be used. Such methods include, without limitation, separation by electrophoresis followed by electroelution, various types of

chromatography (affinity, immunoaffinity, molecular sieve, and/or ion exchange), and/or high pressure liquid chromatography. In some cases, it may be preferable to use more than one of these methods for complete purification.

Human E3 α polypeptides, including fragments, variants, and/or derivatives thereof may also be prepared by chemical synthesis methods (such as solid phase peptide synthesis) using techniques known in the art, such as those set forth by Merrifield *et al.*, *J. Am. Chem. Soc.*, 85 :2149, 1963, Houghten *et al.*, *Proc Natl Acad. Sci. USA*, 82: 5132 1985, and Stewart and Young, *Solid Phase Peptide Synthesis*, Pierce Chemical Co., Rockford, IL (1984). Such polypeptides may be synthesized with or without a methionine on the amino terminus. Chemically synthesized huE3 α polypeptides may be oxidized using methods set forth in these references to form disulfide bridges. Chemically synthesized huE3 α polypeptides are expected to have comparable biological activity to the corresponding huE3 α polypeptides produced recombinantly or purified from natural sources, and thus may be used interchangeably with a recombinant or natural huE3 α polypeptide.

Another means of obtaining a huE3 α polypeptide is via purification from biological samples such as source tissues and/or fluids in which the huE3 α polypeptide is naturally found. Such purification can be conducted using methods for protein purification as described herein. The presence of the huE3 α polypeptide during purification may be monitored using, for example, using an antibody prepared against recombinantly produced huE3 α polypeptide or peptide fragments thereof.

A number of additional methods for producing nucleic acids and polypeptides are known in the art, and the methods can be used to produce polypeptides having specificity for h2520-109. See for example, Roberts *et al.*, *Proc. Natl. Acad. Sci U.S.A.*, 94:12297-12303, 1997, which describes the production of fusion proteins between an mRNA and its encoded peptide. See also Roberts, R., *Curr. Opin. Chem. Biol.*, 3:268-273, 1999. Additionally, U.S. Patent No. 5,824,469 describes methods of obtaining oligonucleotides capable of carrying out a specific biological function. The procedure involves generating a heterogeneous pool of oligonucleotides, each having a 5' randomized sequence, a central preselected sequence, and a 3' randomized sequence. The resulting heterogeneous pool is introduced into a population of cells that do not exhibit the desired biological function. Subpopulations of the cells are then screened for those which exhibit a predetermined biological function. From that subpopulation, oligonucleotides capable of carrying out the desired biological function are isolated.

U.S. Patent Nos. 5,763,192; 5,814,476; 5,723,323; and 5,817,483 describe processes for producing peptides or polypeptides. This is done by producing stochastic genes or fragments thereof, and then introducing these genes into host cells which produce one or more proteins encoded by the stochastic genes. The host cells are then screened to identify those clones producing peptides or polypeptides having the desired activity.

Another method for producing peptides or polypeptides is described in PCT/US98/20094 (WO99/15650) filed by Athersys, Inc. Known as "Random Activation of Gene Expression for Gene Discovery" (RAGE-GD), the process involves the activation of endogenous gene expression or over-expression of a gene by in situ recombination methods. For example, expression of an endogenous gene is activated or increased by integrating a regulatory sequence into the target cell which is capable of activating expression of the gene by non-homologous or illegitimate recombination. The target DNA is first subjected to radiation, and a genetic promoter inserted. The promoter eventually locates a break at the front of a gene, initiating transcription of the gene. This results in expression of the desired peptide or polypeptide.

It will be appreciated that these methods can also be used to create comprehensive IL-17 like protein expression libraries, which can subsequently be used for high throughput phenotypic screening in a variety of assays, such as biochemical assays, cellular assays, and whole organism assays (*e.g.*, plant, mouse, etc.).

Chemical Derivatives

Chemically modified derivatives of the h2520-109 polypeptides may be prepared by one skilled in the art, given the disclosures set forth herein below. h2520-109 polypeptide derivatives are modified in a manner that is different, either in the type or location of the molecules naturally attached to the polypeptide. Derivatives may include molecules formed by the deletion of one or more naturally-attached chemical groups. The polypeptide comprising the amino acid sequence of SEQ ID NO: 2, or a h2520-109 polypeptide variant, may be modified by the covalent attachment of one or more polymers. For example, the polymer selected is typically water soluble so that the protein to which it is attached does not precipitate in an aqueous environment, such as a physiological environment. Included within the scope of suitable polymers is a mixture of polymers. Preferably, for therapeutic use of the end-product preparation, the polymer will be pharmaceutically acceptable.

The polymers each may be of any molecular weight and may be branched or unbranched. The polymers each typically have an average molecular weight of between about 2 kDa to about 100 kDa (the term "about" indicating that in preparations of a water soluble polymer, some molecules will weigh more, some less, than the stated molecular weight). The average molecular weight of each polymer is preferably between about 5 kDa and about 50 kDa, more preferably between about 12 kDa and about 40 kDa and most preferably between about 20 kDa to about 35 kDa. Suitable water soluble polymers or mixtures thereof include, but are not limited to, N-linked or O-linked carbohydrates; sugars; phosphates; polyethylene glycol (PEG) (including the forms of PEG that have been used to derivatize proteins, including mono-(C1-C10) alkoxy- or aryloxy-polyethylene glycol), monomethoxy-polyethylene glycol; dextran (such as low molecular weight dextran of, for example about 6 kD; cellulose, or other carbohydrate-based polymers, poly-(N-vinyl pyrrolidone) polyethylene glycol, propylene glycol homopolymers, a polypropylene oxide/ethylene oxide co-polymer, polyoxyethylated polyols (e.g., glycerol) and polyvinyl alcohol. Also encompassed by the present invention are bifunctional crosslinking molecules which may be used to prepare covalently attached multimers of the polypeptide comprising the amino acid sequence of SEQ ID NO: 2 or a h2520-109 polypeptide variant.

In general, chemical derivatization may be performed under any suitable condition used to react a protein with an activated polymer molecule. Methods for preparing chemical derivatives of polypeptides will generally comprise the steps of (a) reacting the polypeptide with the activated polymer molecule (such as a reactive ester or aldehyde derivative of the polymer molecule) under conditions whereby the polypeptide comprising the amino acid sequence of SEQ ID NO: 2, or a h2520-109 polypeptide variant becomes attached to one or more polymer molecules, and (b) obtaining the reaction product(s). The optimal reaction conditions will be determined based on known parameters and the desired result. For example, the larger the ratio of polymer molecules:protein, the greater the percentage of attached polymer molecule. In one embodiment, the h2520-109 polypeptide derivative may have a single polymer molecule moiety at the amino terminus. (See, for example, U.S. Patent No. 5,234,784). The pegylation of the polypeptide may be specifically carried out by any of the pegylation reactions known in the art, as described for example in the following references: Francis *et al.*, *Focus on Growth Factors*, 3:4-10 (1992); EP 0154316; EP 0401384 and U.S. Patent No. 4,179,337. For example, pegylation may be carried out via

an acylation reaction or an alkylation reaction with a reactive polyethylene glycol molecule (or an analogous reactive water-soluble polymer) as described herein. For the acylation reactions, the polymer(s) selected should have a single reactive ester group. For reductive alkylation, the polymer(s) selected should have a single reactive aldehyde group. A reactive aldehyde is, for example, polyethylene glycol propionaldehyde, which is water stable, or mono C1-C10 alkoxy or aryloxy derivatives thereof (see U.S. Patent No. 5,252,714).

In another embodiment, h2520-109 polypeptides may be chemically coupled to biotin, and the biotin/h2520-109 polypeptide molecules which are conjugated are then allowed to bind to avidin, resulting in tetravalent avidin/biotin/h2520-109 polypeptide molecules. h2520-109 polypeptides may also be covalently coupled to dinitrophenol (DNP) or trinitrophenol (TNP) and the resulting conjugates precipitated with anti-DNP or anti-TNP-IgM to form decameric conjugates with a valency of 10.

Generally, conditions which may be alleviated or modulated by the administration of the present h2520-109 polypeptide derivatives include those described herein for h2520-109 polypeptides. However, the h2520-109 polypeptide derivatives disclosed herein may have additional activities, enhanced or reduced biological activity, or other characteristics, such as increased or decreased half-life, as compared to the non-derivatized molecules.

Selective Binding Agents

As used herein, the term "selective binding agent" refers to a molecule which has specificity for one or more huE3 α polypeptides. Suitable selective binding agents include, but are not limited to, antibodies and derivatives thereof, polypeptides, and small molecules. Suitable selective binding agents may be prepared using methods known in the art. An exemplary huE3 α polypeptide selective binding agent of the present invention is capable of binding a certain portion of the huE3 α polypeptide thereby inhibiting the binding of a cofactor to the huE3 α polypeptide.

Human E3 α polypeptides, fragments, variants, and derivatives may be used to prepare selective binding agents (such as antibodies) using methods known in the art; wherein huE3 α polypeptide refers to either huE3 α I or huE3 α II polypeptide. Thus, selective binding agents such as antibodies and antibody fragments that bind huE3 α polypeptides are within the scope of the present invention. The antibodies may

be polyclonal, monospecific polyclonal, monoclonal, recombinant, chimeric, humanized, human, single chain, and/or bispecific.

5 Polyclonal antibodies directed toward a huE3 α polypeptide generally are raised in animals (*e.g.*, rabbits or mice) by multiple subcutaneous or intraperitoneal injections of huE3 α polypeptide and an adjuvant. It may be useful to conjugate a huE3 α polypeptide, or a variant, fragment, or derivative thereof to a carrier protein that is immunogenic in the species to be immunized, such as keyhole limpet heocyanin, serum, albumin, bovine thyroglobulin, or soybean trypsin inhibitor. Also, aggregating agents such as alum are used to enhance the immune response. After immunization, the animals
10 are bled and the serum is assayed for anti-huE3 α antibody titer.

Monoclonal antibodies directed toward huE3 α polypeptides are produced using any method which provides for the production of antibody molecules by continuous cell lines in culture. Examples of suitable methods for preparing monoclonal antibodies include the hybridoma methods of Kohler *et al.*, *Nature*, 256: 495-497, 1975
15 and the human B-cell hybridoma method, Kozbor, *J. Immunol.*, 133: 3001, 1984; Brodeur *et al.*, *Monoclonal Antibody Production Techniques and Applications*, pp. 51-63 (Marcel Dekker, Inc., New York, 1987). Also provided by the invention are hybridoma cell lines which produce monoclonal antibodies reactive with huE3 α polypeptides.

Monoclonal antibodies of the invention may be modified for use as
20 therapeutics. One embodiment is a "chimeric" antibody in which a portion of the heavy and/or light chain is identical with or homologous to a corresponding sequence in antibodies derived from a particular species or belonging to a particular antibody class or subclass, while the remainder of the chain(s) is identical with or homologous to a corresponding sequence in antibodies derived from another species or belonging to
25 another antibody class or subclass. Also included are fragments of such antibodies, so long as they exhibit the desired biological activity. See, U.S. Patent No. 4,816,567; Morrison *et al.*, *Proc. Natl. Acad. Sci.*, 81: 6851-6855 (1985).

In another embodiment, a monoclonal antibody of the invention is a
"humanized" antibody. Methods for humanizing non-human antibodies are well known
30 in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source which is non-human. Humanization can be performed following methods known in the art (Jones *et al.*, *Nature* 321: 522-525, 1986; Riechmann *et al.*, *Nature*, 332: 323-327 (1988); Verhoeyen *et al.*, *Science* 239:1534-

1536, 1988), by substituting rodent complementarity-determining regions (CDRs) for the corresponding regions of a human antibody.

Also encompassed by the invention are human antibodies which bind huE3 α polypeptides, fragments, variants and/or derivatives. Such antibodies are produced by immunization with a huE3 α antigen (*i.e.*, having at least 6 contiguous amino acids), optionally conjugated to a carrier, of transgenic animals (*e.g.*, mice) that are capable of producing a repertoire of human antibodies in the absence of endogenous immunoglobulin production. *See*, for example, Jakobovits *et al.*, *Proc. Natl. Acad. Sci.*, 90: 2551-2555, 1993; Jakobovits *et al.*, *Nature* 362: 255-258, 1993; Bruggermann *et al.*, *Year in Immuno.*, 7: 33 (1993). In one method, such transgenic animals are produced by incapacitating the endogenous loci encoding the heavy and light immunoglobulin chains therein, and inserting loci encoding human heavy and light chain proteins into the genome thereof. Partially modified animals, that is those having less than the full complement of modifications, are then cross-bred to obtain an animal having all of the desired immune system modifications. When administered an immunogen, these transgenic animals produce antibodies with human variable regions, including human (rather than *e.g.*, murine) antibodies which are immunospecific for these antigens. *See* PCT application Nos. PCT/US96/05928 and PCT/US93/06926. Additional methods are described in U.S. Patent No. 5,545,807, PCT application Nos. PCT/US91/245, PCT/GB89/01207, and in EP 546073B1 and EP 546073A1.

Human antibodies can also be produced from phage-display libraries (Hoogenboom *et al.*, *J. Mol. Biol.* 227: 381 1991; Marks *et al.*, *J. Mol. Biol.* 222: 581, 1991). These processes mimic immune selection through the display of antibody repertoires on the surface of filamentous bacteriophage, and subsequent selection of phage by their binding to an antigen of choice. One such technique is described in PCT Application WO99/10494, filed in the name of Adams *et al.*, which describes the isolation of high affinity and functional agonistic antibodies for MPL- and msk-receptors using such an approach.

Chimeric, CDR grafted, and humanized antibodies are typically produced by recombinant methods. Nucleic acids encoding the antibodies are introduced into host cells and expressed using materials and procedures described herein. In a preferred embodiment, the antibodies are produced in mammalian host cells, such as CHO cells. Human antibodies may be produced by the expression of recombinant DNA in host cells or by expression in hybridoma cells as described herein.

For diagnostic applications, in certain embodiments, anti-huE3 α antibodies typically will be labeled with a detectable moiety. The detectable moiety can be any one which is capable of producing, either directly or indirectly, a detectable signal. For example, the detectable moiety may be a radioisotope, such as ^3H , ^{14}C , ^{32}P , ^{35}S , or ^{125}I , a fluorescent or chemiluminescent compound, such as fluorescein isothiocyanate, rhodamine, or luciferin; or an enzyme, such as alkaline phosphatase, β -galactosidase, or horseradish peroxidase (Bayer *et al.*, *Meth. Enz.*, 184: 138-163 1990).

The anti-huE3 α antibodies of the invention may be employed in any known assay method, such as competitive binding assays, direct and indirect sandwich assays, and immunoprecipitation assays (Sola, *Monoclonal Antibodies: A Manual of Techniques*, pp. 147-158 (CRC Press, Inc., 1987)) for the detection and quantitation of huE3 α polypeptides. The antibodies will bind huE3 α polypeptides with an affinity which is appropriate for the assay method being employed.

Competitive binding assays rely on the ability of a labeled standard (*e.g.*, a huE3 α polypeptide, or an immunologically reactive portion thereof) to compete with the test sample analyte (a huE3 α polypeptide) for binding with a limited amount of anti-huE3 α antibody. The amount of a huE3 α polypeptide in the test sample is inversely proportional to the amount of standard that becomes bound to the antibodies. To facilitate determining the amount of standard that becomes bound, the antibodies typically are insolubilized before or after the competition, so that the standard and analyte that are bound to the antibodies may conveniently be separated from the standard and analyte which remain unbound.

Sandwich assays typically involve the use of two antibodies, each capable of binding to a different immunogenic portion, or epitope, of the protein to be detected and/or quantitated. In a sandwich assay, the test sample analyte is typically bound by a first antibody which is immobilized on a solid support, and thereafter a second antibody binds to the analyte, thus forming an insoluble three part complex. *See, e.g.*, U.S. Patent No. 4,376,110. The second antibody may itself be labeled with a detectable moiety (direct sandwich assays) or may be measured using an anti-immunoglobulin antibody that is labeled with a detectable moiety (indirect sandwich assays). For example, one type of sandwich assay is an enzyme-linked immunosorbent assay (ELISA), in which case the detectable moiety is an enzyme.

The selective binding agents, including anti-h2520-109 antibodies, are also useful for *in vivo* imaging. An antibody labeled with a detectable moiety may be

administered to an animal, preferably into the bloodstream, and the presence and location of the labeled antibody in the host is assayed. The antibody may be labeled with any moiety that is detectable in an animal, whether by nuclear magnetic resonance, radiology, or other detection means known in the art.

5 Selective binding agents of the invention, including anti-h2520-109 antibodies, may be used as therapeutics. These therapeutic agents are generally agonists or antagonists, in that they either enhance or reduce, respectively, at least one of the biological activities of a h2520-109 polypeptide. In one embodiment, antagonist antibodies of the invention are antibodies or binding fragments thereof which are capable of specifically binding to a h2520-109 polypeptide and which are capable of inhibiting or eliminating the functional activity of a h2520-109 polypeptide *in vivo* or *in vitro*. In preferred embodiments, the selective binding agent, *e.g.*, an antagonist antibody will inhibit the functional activity of a h2520-109 polypeptide by at least about 50%, and preferably by at least about 80%. In another embodiment, the selective binding agent 10 may be an antibody that is capable of interacting with a h2520-109 binding partner (a ligand, co-factor, or receptor) thereby inhibiting or eliminating h2520-109 activity *in vitro* or *in vivo*. Selective binding agents, including agonist and antagonist anti-h2520-109 antibodies are identified by screening assays which are well known in the art.

The invention also relates to a kit comprising huE3 α selective binding agents (such as antibodies) and other reagents useful for detecting huE3 α levels in biological samples. Such reagents may include a secondary activity, a detectable label, blocking serum, positive and negative control samples, and detection reagents. 20

Genetically Engineered Non-Human Animals

25 Additionally included within the scope of the present invention are non-human animals such as mice, rats, or other rodents, rabbits, goats, or sheep, or other farm animals, in which the gene (or genes) encoding a native E3 α ubiquitin ligase polypeptide (such as E3 α I or E3 α II) has (have) been disrupted ("knocked out") such that the level of expression of this gene or genes is (are) significantly decreased or completely 30 abolished. Such animals may be prepared using techniques and methods such as those described in U.S. Patent No. 5,557,032.

The present invention further includes non-human animals such as mice, rats, or other rodents, rabbits, goats, or sheep, or other farm animals, in which either the native form of the E3 α ubiquitin ligase polypeptide gene(s) for that animal or a

heterologous E3 α ubiquitin ligase polypeptide gene(s) is (are) over expressed by the animal, thereby creating a "transgenic" animal. Such transgenic animals may be prepared using well known methods such as those described in U.S. Patent No 5,489,743 and PCT application No. WO94/28122.

5 The present invention further includes non-human animals in which the promoter for one or more of the E3 α ubiquitin ligase polypeptides of the present invention is either activated or inactivated (*e.g.*, by using homologous recombination methods) to alter the level of expression of one or more of the native E3 α ubiquitin ligase polypeptides.

10 These non-human animals may be used for drug candidate screening. In such screening, the impact of a drug candidate on the animal may be measured. For example, drug candidates may decrease or increase the expression of the E3 α ubiquitin ligase polypeptide gene. In certain embodiments, the amount of E3 α ubiquitin ligase polypeptide, or a fragment(s), that is produced may be measured after the exposure of
15 the animal to the drug candidate. Additionally, in certain embodiments, one may detect the actual impact of the drug candidate on the animal. For example, the overexpression of a particular gene may result in, or be associated with, a disease or pathological condition. In such cases, one may test a drug candidate's ability to decrease expression of the gene or its ability to prevent or inhibit a pathological condition. In other
20 examples, the production of a particular metabolic product such as a fragment of a polypeptide, may result in, or be associated with, a disease or pathological condition. In such cases, one may test a drug candidate's ability to decrease the production of such a metabolic product or its ability to prevent or inhibit a pathological condition.

25 Microarray

It will be appreciated that DNA microarray technology can be utilized in accordance with the present invention. DNA microarrays are miniature, high density arrays of nucleic acids positioned on a solid support, such as glass. Each cell or element within the array has numerous copies of a single species of DNA which acts as a target
30 for hybridization for its cognate mRNA. In expression profiling using DNA microarray technology, mRNA is first extracted from a cell or tissue sample and then converted enzymatically to fluorescently labeled cDNA. This material is hybridized to the microarray and unbound cDNA is removed by washing. The expression of discrete genes represented on the array is then visualized by quantitating the amount of labeled

cDNA which is specifically bound to each target DNA. In this way, the expression of thousands of genes can be quantitated in a high throughput, parallel manner from a single sample of biological material.

This high throughput expression profiling has a broad range of applications with respect to the h2520-109 molecules of the invention, including, but not limited to: the identification and validation of h2520-109 disease-related genes as targets for therapeutics; molecular toxicology of h2520-109 molecules and inhibitors thereof; stratification of populations and generation of surrogate markers for clinical trials; and the enhancement of an h2520-109-related small molecule drug discovery by aiding in the identification of selective compounds in high throughput screens (HTS).

Assaying for other modulators of huE3 α polypeptide activity

In some situations, it may be desirable to identify molecules that are modulators, *i.e.*, antagonists and agonists, of the activity of huE3 α polypeptide.

Natural or synthetic molecules that modulate huE3 α polypeptides can be identified using one or more screening assays, such as those described herein. Such molecules may be administered either in an *ex vivo* manner, or in an *in vivo* manner by injection, or by oral delivery, implantation device, or the like.

The following definition is used herein for describing the assays. "Test molecule(s)" refers to the molecule(s) that is/are under evaluation for the ability to modulate (*i.e.*, increase or decrease) the activity of a huE3 α polypeptide. Most commonly, a test molecule will interact directly with a huE3 α polypeptide. However, it is also contemplated that a test molecule may also modulate huE3 α polypeptide activity indirectly, such as by affecting huE3 α gene expression, or by binding to a huE3 α binding partner (*e.g.*, receptor, co-factor or ligand). In one embodiment, a test molecule will bind to a huE3 α polypeptide with an affinity constant of at least about 10^{-6} M, preferably about 10^{-8} M, more preferably about 10^{-9} M, and even more preferably about 10^{-10} M.

Methods for identifying compounds which interact with huE3 α polypeptides are encompassed by the present invention. In certain embodiments, a huE3 α polypeptide is incubated with a test molecule under conditions which permit the interaction of the test molecule with a huE3 α polypeptide, and the extent of the interaction can be measured. The test molecule(s) can be screened in a substantially purified form or in a crude mixture. Test molecule(s) can be nucleic acid molecules,

proteins, peptides, carbohydrates, lipids, or small molecular weight organic or inorganic compounds. Once a set of has been identified as interacting with a huE3 α polypeptide, the molecules may be further evaluated for their ability to increase or decrease huE3 α activity.

5 The measurement of the interaction of test molecules with huE3 α polypeptides may be carried out in several formats, including cell-based binding assays, membrane binding assays, solution-phase assays and immunoassays. In general, test molecules are incubated with a huE3 α polypeptide for a specified period of time, and huE3 α activity is determined by one or more assays described herein for measuring
10 biological activity.

 The interaction of test molecules with huE3 α polypeptides may also be assayed directly using polyclonal or monoclonal antibodies in an immunoassay. Alternatively, modified forms of huE3 α polypeptides containing epitope tags as described herein may be used in solution and immunoassays.

15 In certain embodiments, a huE3 α polypeptide agonist or antagonist may be a protein, peptide, carbohydrate, lipid, or small molecular weight molecule which interacts with huE3 α polypeptide to regulate its activity. Molecules which regulate huE3 α polypeptide expression include nucleic acids which are complementary to nucleic acids encoding a huE3 α polypeptide, or are complementary to nucleic acids sequences
20 which direct or control the expression of huE3 α polypeptide, and which act as antisense regulators of expression.

 Once a set of test molecules has been identified as interacting with a h2520-109 polypeptide, the molecules may be further evaluated for their ability to increase or decrease h2520-109 polypeptide activity. The measurement of the
25 interaction of test molecules with h2520-109 polypeptides may be carried out in several formats, including cell-based binding assays, membrane binding assays, solution-phase assays and immunoassays. In general, test molecules are incubated with a h2520-109 polypeptide for a specified period of time, and h2520-109 polypeptide activity is determined by one or more assays for measuring biological activity.

30 The interaction of test molecules with h2520-109 polypeptides may also be assayed directly using polyclonal or monoclonal antibodies in an immunoassay. Alternatively, modified forms of h2520-109 polypeptides containing epitope tags as described herein may be used in immunoassays.

In the event that h2520-109 polypeptides display biological activity through an interaction with a binding partner (*e.g.*, a receptor, a ligand or a co-factor), a variety of *in vitro* assays may be used to measure the binding of a h2520-109 polypeptide to the corresponding binding partner (such as a selective binding agent, receptor, ligand, or co-factor). These assays may be used to screen test molecules for their ability to increase or decrease the rate and/or the extent of binding of a h2520-109 polypeptide to its binding partner. In one assay, a h2520-109 polypeptide is immobilized in the wells of a microtiter plate. Radiolabeled h2520-109 binding partner (for example, iodinated h2520-109 binding partner) and the test molecule(s) can then be added either one at a time (in either order) or simultaneously to the wells. After incubation, the wells can be washed and counted using a scintillation counter, for radioactivity to determine the extent to which the binding partner bound to h2520-109 polypeptide. Typically, the molecules will be tested over a range of concentrations, and a series of control wells lacking one or more elements of the test assays can be used for accuracy in the evaluation of the results. An alternative to this method involves reversing the "positions" of the proteins, *i.e.*, immobilizing h2520-109 binding partner to the microtiter plate wells, incubating with the test molecule and radiolabeled h2520-109 polypeptide, and determining the extent of h2520-109 polypeptide binding. *See*, for example, chapter 18, *Current Protocols in Molecular Biology*, Ausubel *et al.*, eds., John Wiley & Sons, New York, NY (1995).

As an alternative to radiolabelling, a h2520-109 polypeptide or its binding partner may be conjugated to biotin and the presence of biotinylated protein can then be detected using streptavidin linked to an enzyme, such as horseradish peroxidase (HRP) or alkaline phosphatase (AP), that can be detected colorometrically, or by fluorescent tagging of streptavidin. An antibody directed to a h2520-109 polypeptide or to a h2520-109 binding partner and conjugated to biotin may also be used and can be detected after incubation with enzyme-linked streptavidin linked to AP or HRP.

A h2520-109 polypeptide or a h2520-109 like binding partner can also be immobilized by attachment to agarose beads, acrylic beads or other types of such inert solid phase substrates. The substrate-protein complex can be placed in a solution containing the complementary protein and the test compound. After incubation, the beads can be precipitated by centrifugation, and the amount of binding between a h2520-109 polypeptide and its binding partner can be assessed using the methods described herein. Alternatively, the substrate-protein complex can be immobilized in a column,

and the test molecule and complementary protein are passed through the column. The formation of a complex between a h2520-109 polypeptide and its binding partner can then be assessed using any of the techniques set forth herein, *i.e.*, radiolabelling, antibody binding or the like.

5 Another *in vitro* assay that is useful for identifying a test molecule which increases or decreases the formation of a complex between a h2520-109 polypeptide and a h2520-109 binding partner is a surface plasmon resonance detector system such as the BIAcore assay system (Pharmacia, Piscataway, NJ). The BIAcore system may be carried out using the manufacturer's protocol. This assay essentially involves the
10 covalent binding of either h2520-109 polypeptide or a h2520-109 binding partner to a dextran-coated sensor chip which is located in a detector. The test compound and the other complementary protein can then be injected, either simultaneously or sequentially, into the chamber containing the sensor chip. The amount of complementary protein that binds can be assessed based on the change in molecular mass which is physically
15 associated with the dextran-coated side of the sensor chip; the change in molecular mass can be measured by the detector system.

In some cases, it may be desirable to evaluate two or more test compounds together for their ability to increase or decrease the formation of a complex between a h2520-109 polypeptide and a h2520-109 binding partner. In these cases, the assays set
20 forth herein can be readily modified by adding such additional test compound(s) either simultaneous with, or subsequent to, the first test compound. The remainder of the steps in the assay are set forth herein .

In vitro assays such as those described herein may be used advantageously to screen large numbers of compounds for effects on complex formation by h2520-109
25 polypeptide and h2520-109 binding partner. The assays may be automated to screen compounds generated in phage display, synthetic peptide, and chemical synthesis libraries.

Compounds which increase or decrease the formation of a complex between a h2520-109 polypeptide and a h2520-109 binding partner may also be screened in cell
30 culture using cells and cell lines expressing either h2520-109 polypeptide or h2520-109 binding partner. Cells and cell lines may be obtained from any mammal, but preferably will be from human or other primate, canine, or rodent sources. The binding of a h2520-109 polypeptide to cells expressing h2520-109 binding partner at the surface is evaluated in the presence or absence of test molecules, and the extent of binding may be

determined by, for example, flow cytometry using a biotinylated antibody to a h2520-109 binding partner. Cell culture assays can be used advantageously to further evaluate compounds that score positive in protein binding assays described herein

Cell cultures can be used to screen the impact of a drug candidate. For example, drug candidates may decrease or increase the expression of the huE3 α polypeptide gene. In certain embodiments, the amount of huE3 α polypeptide or a fragment(s) that is produced may be measured after exposure of the cell culture to the drug candidate. In certain embodiments, one may detect the actual impact of the drug candidate on the cell culture. For example, the overexpression of a particular gene may have a particular impact on the cell culture. In such cases, one may test a drug candidate's ability to increase or decrease the expression of the gene or its ability to prevent or inhibit a particular impact on the cell culture. In other examples, the production of a particular metabolic product such as a fragment of a polypeptide, may result in, or be associated with, a disease or pathological condition. In such cases, one may test a drug candidate's ability to decrease the production of such a metabolic product in a cell culture.

A yeast two hybrid system (Chien *et al.*, *Proc. Natl. Acad. Sci. USA*, 88: 9578-9583, 1991) can be used to identify novel polypeptides that bind to a yeast-two hybrid bait construct can be generated in a vector (such as the pAS2-1 from Clontech) which encodes a yeast-two hybrid domain fused to the huE3 α polynucleotide. This bait construct may be used to screen human cDNA libraries wherein the cDNA library sequences are fused to GAL4 activation domains. Positive interactions will result in the activation of a reporter gene such as β -gal. Positive clones emerging from the screening may be characterized further to identify interacting proteins.

Internalizing Proteins

The TAT protein sequence (from HIV) can be used to internalize proteins into a cell by targeting the lipid bi-layer component of the cell membrane. See *e.g.*, Falwell *et al.*, *Proc. Natl. Acad. Sci.*, 91: 664-668, 1994. For example, an 11 amino acid sequence (YGRKKRRQRRR; SEQ ID NO: 16) of the HIV TAT protein (termed the "protein transduction domain", or TAT PDT) has been shown to mediate delivery of large bioactive proteins such as β -galactosidase and p27Kip across the cytoplasmic membrane and the nuclear membrane of a cell. See Schwarze *et al.*, *Science*, 285: 1569-1572, 1999; and Nagahara *et al.*, *Nature Medicine*, 4: 1449-1452, 1998. Schwartz *et al.* (*Science*, 285: 1569-72, 1999) demonstrated that cultured cells acquired β -gal

activity when exposed to a fusion of the TAT PDT and β -galactosidase. Injection of mice with the TAT- β -gal fusion proteins resulted in β -gal expression in a number of tissues, including liver, kidney, lung, heart, and brain tissue.

It will thus be appreciated that the TAT protein sequence may be used to internalize a desired protein or polypeptide into a cell. In the context of the present invention, the TAT protein sequence can be fused to another molecule such as a huE3 α antagonist (*i.e.*: anti-huE3 α selective binding agent or small molecule) and administered intracellularly to inhibit the activity of the huE3 α molecule. Where desired, the huE3 α protein itself, or a peptide fragment or modified form of huE3 α , may be fused to such a protein transducer for administering to cells using the procedures, described above.

Therapeutic Uses

The huE3 α nucleic acid molecules, polypeptides, and antagonists thereof (including, but not limited to, anti-huE3 α selective binding agents) can be used to treat, diagnose, and/or prevent a number of diseases, conditions, and disorders, including but not limited to cachexia, muscle wasting diseases and other catabolic disorders such as cancer cachexia, renal cachexia, inflammatory cachexia, muscle wasting disorders associated with metabolic acidosis, uremia, burns, hyperthyroidism, Cushing's syndrome and fasting, and denervation atrophy, diabetes mellitus, sepsis and AIDS wasting syndrome.

Those skilled in the art will recognize that many combinations of deletions, insertions, and substitutions (individually or collectively "variant(s)" herein) can be made within the amino acid sequences of the huE3 α polypeptide, provided that the resulting molecule is biologically active (*e.g.*, possesses the ability to affect one or more of the diseases and disorders such as those recited herein).

As contemplated by the present invention, a polypeptide, or antagonist thereof (including, but not limited to, anti-huE3 α selective binding agents) may be administered as an adjunct to other therapy and also with other pharmaceutical compositions suitable for the indication being treated. A polypeptide and any of one or more additional therapies or pharmaceutical formulations may be administered separately, sequentially, or simultaneously.

In a specific embodiment, the present invention is directed to the use of a huE3 α polypeptide, or antagonist (including, but not limited to, anti-huE3 α selective binding agents) thereof in combination (pretreatment, post-treatment, or concurrent

treatment) with secreted or soluble human fas antigen or recombinant versions thereof (WO96/20206 and Mountz *et al.*, *J. Immunology*, 155: 4829-4837; and EP 510 691. WO96/20206 discloses secreted human fas antigen (native and recombinant, including an Ig fusion protein), methods for isolating the genes responsible for coding the soluble recombinant human fas antigen, methods for cloning the gene in suitable vectors and cell types, and methods for expressing the gene to produce the inhibitors. EP 510 691 teaches DNAs coding for human fas antigen, including soluble fas antigen, vectors expressing for said DNAs and transformants transfected with the vector. When administered parenterally, doses of a secreted or soluble fas antigen fusion protein each are generally from about 1 microgram/kg to about 100 micrograms/kg.

Treatment of the diseases and disorders recited herein can include the use of first line drugs for control of pain and inflammation; these drugs are classified as non-steroidal, anti-inflammatory drugs (NSAIDs). Secondary treatments include corticosteroids, slow acting antirheumatic drugs (SAARDs), or disease modifying (DM) drugs. Information regarding the following compounds can be found in The Merck Manual of Diagnosis and Therapy, Sixteenth Edition, Merck, Sharp & Dohme Research Laboratories, Merck & Co., Rahway, NJ (1992) and in Pharmaprojects, PJB Publications Ltd.

In a specific embodiment, the present invention is directed to the use of a huE3 α , or antagonist (including, but not limited to, anti-huE3 α selective binding agents) and any of one or more NSAIDs for the treatment of the diseases and disorders recited herein. NSAIDs owe their anti-inflammatory action, at least in part, to the inhibition of prostaglandin synthesis (Goodman and Gilman in "The Pharmacological Basis of Therapeutics," MacMillan 7th Edition (1985)). NSAIDs can be characterized into at least nine groups: (1) salicylic acid derivatives; (2) propionic acid derivatives; (3) acetic acid derivatives; (4) fenamic acid derivatives; (5) carboxylic acid derivatives; (6) butyric acid derivatives; (7) oxicams; (8) pyrazoles and (9) pyrazolones.

In another specific embodiment, the present invention is directed to the use of an huE3 α polypeptide, or antagonist (including, but not limited to, anti-huE3 α selective binding agents) in combination (pretreatment, post-treatment, or concurrent treatment) with any of one or more salicylic acid derivatives, prodrug esters or pharmaceutically acceptable salts thereof. Such salicylic acid derivatives, prodrug esters and pharmaceutically acceptable salts thereof comprise: acetaminosalol, aloxiprin, aspirin, benorylate, bromosaligenin, calcium acetylsalicylate, choline magnesium

trisalicylate, magnesium salicylate, choline salicylate, diflusal, etersalate, fendosal, gentisic acid, glycol salicylate, imidazole salicylate, lysine acetylsalicylate, mesalamine, morpholine salicylate, 1-naphthyl salicylate, olsalazine, parsalimide, phenyl acetylsalicylate, phenyl salicylate, salacetamide, salicylamide O-acetic acid, salsalate, sodium salicylate and sulfasalazine. Structurally related salicylic acid derivatives having similar analgesic and anti-inflammatory properties are also intended to be encompassed by this group.

In an additional specific embodiment, the present invention is directed to the use of an huE3 α polypeptide, or antagonist (including, but not limited to, anti-huE3 α selective binding agents) in combination (pretreatment, post-treatment, or concurrent treatment) with any of one or more propionic acid derivatives, prodrug esters or pharmaceutically acceptable salts thereof. The propionic acid derivatives, prodrug esters, and pharmaceutically acceptable salts thereof comprise: alminoprofen, benoxaprofen, bucloxic acid, carprofen, dexindoprofen, fenoprofen, flunoxaprofen, fluprofen, flurbiprofen, furocloprofen, ibuprofen, ibuprofen aluminum, ibuprofen, indoprofen, isoprofen, ketoprofen, loxoprofen, miroprofen, naproxen, naproxen sodium, oxaprozin, piketoprofen, pimeprofen, pirprofen, pranoprofen, protizinic acid, pyridoxiprofen, suprofen, tiaprofenic acid and tioxaprofen. Structurally related propionic acid derivatives having similar analgesic and anti-inflammatory properties are also intended to be encompassed by this group.

In yet another specific embodiment, the present invention is directed to the use of a huE3 α polypeptide, or antagonist (including, but not limited to, anti-huE3 α selective binding agents) in combination (pretreatment, post-treatment, or concurrent treatment) with any of one or more acetic acid derivatives, prodrug esters or pharmaceutically acceptable salts thereof. The acetic acid derivatives, prodrug esters, and pharmaceutically acceptable salts thereof comprise: acetaminophen, alclometacin, alclometacin, amfenac, buprenorphine, cinmetacin, clopirac, delmetacin, diclofenac potassium, diclofenac sodium, etodolac, felbinac, fenclofenac, fenclozic acid, fentiazac, furofenac, glucametacin, ibufenac, indomethacin, isofezolac, isoxepac, lonazolac, metiazinic acid, oxametacin, oxpinac, pimetacin, proglumetacin, sulindac, talmetacin, tiaramide, tiopinac, tolmetin, tolmetin sodium, zidometacin and zomepirac. Structurally related acetic acid derivatives having similar analgesic and anti-inflammatory properties are also intended to be encompassed by this group.

In another specific embodiment, the present invention is directed to the use of a huE3 α polypeptide, or antagonist (including, but not limited to, anti-huE3 α selective binding agents) in combination (pretreatment, post-treatment, or concurrent treatment) with any of one or more fenamic acid derivatives, prodrug esters or pharmaceutically acceptable salts thereof. The fenamic acid derivatives, prodrug esters and pharmaceutically acceptable salts thereof comprise: enfenamic acid, etofenamate, flufenamic acid, isonixin, meclofenamic acid, meclofenamate sodium, medofenamic acid, mefenamic acid, niflumic acid, talniflumate, terofenamate, tolfenamic acid and ufenamate. Structurally related fenamic acid derivatives having similar analgesic and anti-inflammatory properties are also intended to be encompassed by this group.

In an additional specific embodiment, the present invention is directed to the use of a huE3 α polypeptide, or antagonist (including, but not limited to, anti-huE3 α selective binding agents) in combination (pretreatment, post-treatment, or concurrent treatment) with any of one or more carboxylic acid derivatives, prodrug esters or pharmaceutically acceptable salts thereof. The carboxylic acid derivatives, prodrug esters, and pharmaceutically acceptable salts thereof which can be used comprise: clidanac, diflunisal, flufenisal, inoridine, ketorolac and tinoridine. Structurally related carboxylic acid derivatives having similar analgesic and anti-inflammatory properties are also intended to be encompassed by this group.

In yet another specific embodiment, the present invention is directed to the use of a huE3 α polypeptide, or antagonist (including, but not limited to, anti-huE3 α selective binding agents) in combination (pretreatment, post-treatment, or concurrent treatment) with any of one or more butyric acid derivatives, prodrug esters or pharmaceutically acceptable salts thereof. The butyric acid derivatives, prodrug esters, and pharmaceutically acceptable salts thereof comprise: bumadizon, butibufen, fenbufen and xenbucin. Structurally related butyric acid derivatives having similar analgesic and anti-inflammatory properties are also intended to be encompassed by this group.

In another specific embodiment, the present invention is directed to the use of a huE3 α polypeptide, or antagonist (including, but not limited to, anti-huE3 α selective binding agents) in combination (pretreatment, post-treatment, or concurrent treatment) with any of one or more oxicams, prodrug esters, or pharmaceutically acceptable salts thereof. The oxicams, prodrug esters, and pharmaceutically acceptable salts thereof comprise: droxicam, enolicam, isoxicam, piroxicam, sudoxicam, tenoxicam and 4-hydroxyl-1,2-benzothiazine 1,1-dioxide 4-(N-phenyl)-carboxamide. Structurally

related oxicams having similar analgesic and anti-inflammatory properties are also intended to be encompassed by this group.

In still another specific embodiment, the present invention is directed to the use of a huE3 α polypeptide, or antagonist (including, but not limited to, anti-huE3 α selective binding agents) in combination (pretreatment, post-treatment, or concurrent treatment) with any of one or more pyrazoles, prodrug esters, or pharmaceutically acceptable salts thereof. The pyrazoles, prodrug esters, and pharmaceutically acceptable salts thereof which may be used comprise: difenamizole and epirizole. Structurally related pyrazoles having similar analgesic and anti-inflammatory properties are also intended to be encompassed by this group.

In an additional specific embodiment, the present invention is directed to the use of a huE3 α polypeptide, or antagonist (including, but not limited to, anti-huE3 α selective binding agents) in combination (pretreatment, post-treatment or, concurrent treatment) with any of one or more pyrazolones, prodrug esters, or pharmaceutically acceptable salts thereof. The pyrazolones, prodrug esters and pharmaceutically acceptable salts thereof which may be used comprise: apazone, azapropazone, benzpiperylon, feprazone, mofebutazone, morazone, oxyphenbutazone, phenylbutazone, pipebuzone, propylphenazone, ramifenazone, suxibuzone and thiazolinobutazone. Structurally related pyrazalones having similar analgesic and anti-inflammatory properties are also intended to be encompassed by this group.

In another specific embodiment, the present invention is directed to the use of a huE3 α polypeptide, or antagonist (including, but not limited to, anti-huE3 α selective binding agents) in combination (pretreatment, post-treatment, or concurrent treatment) with any of one or more of the following NSAIDs: e-acetamidocaproic acid, S-adenosyl-methionine, 3-amino-4-hydroxybutyric acid, amixetrine, anitrazafen, antrafenine, bendazac, bendazac lysinate, benzydamine, beprozin, broperamole, bucolome, bufezolac, ciproquazone, cloximate, dazidamine, deboxamet, detomidine, difenpiramide, difenpyramide, difisalamine, ditazol, emorfazone, fanetizole mesylate, fenflumizole, floctafenine, flumizole, flunixin, fluproquazone, fopirtoline, fosfosal, guaimesal, guaiazolene, isonixirn, lefetamine HCl, leflunomide, lofemizole, lotifazole, lysin clonixinate, meseclazone, nabumetone, nictindole, nimesulide, orgotein, orpanoxin, oxaceprol, oxapadol, paranyline, perisoxal, perisoxal citrate, pifoxime, piproxen, pirazolac, pirfenidone, proquazone, proxazole, thielavin B, tiflamizole, timegadine, tolectin, tolpadol, tryptamid and those designated by company code number such as

480156S, AA861, AD1590, AFP802, AFP860, AI77B, AP504, AU8001, BPPC, BW540C, CHINOIN 127, CN100, EB382, EL508, F1044, FK-506, GV3658, ITF182, KCNTEI6090, KME4, LA2851, MR714, MR897, MY309, ONO3144, PR823, PV102, PV108, R830, RS2131, SCR152, SH440, SIR133, SPAS510, SQ27239, ST281, SY6001, TA60, TAI-901 (4-benzoyl-1-indancarboxylic acid), TVX2706, U60257, UR2301 and WY41770. Structurally related NSAIDs having similar analgesic and anti-inflammatory properties to the NSAIDs are also intended to be encompassed by this group.

In still another specific embodiment, the present invention is directed to the use of a huE3 α polypeptide, or antagonist (including, but not limited to, anti-huE3 α selective binding agents) in combination (pretreatment, post-treatment or concurrent treatment) with any of one or more corticosteroids, prodrug esters or pharmaceutically acceptable salts thereof for the treatment of the diseases and disorders recited herein, including acute and chronic inflammation such as rheumatic diseases, graft versus host disease and multiple sclerosis. Corticosteroids, prodrug esters and pharmaceutically acceptable salts thereof include hydrocortisone and compounds which are derived from hydrocortisone, such as 21-acetoxypregnenolone, alclomerasone, algestone, amcinonide, beclomethasone, betamethasone, betamethasone valerate, budesonide, chloroprednisone, clobetasol, clobetasol propionate, clobetasone, clobetasone butyrate, clocortolone, cloprednol, corticosterone, cortisone, cortivazol, deflazacon, desonide, desoximetasone, dexamethasone, diflorasone, diflucortolone, difluprednate, enoxolone, fluazacort, flucoronide, flumethasone, flumethasone pivalate, flucinolone acetonide, flunisolide, fluocinonide, fluorocinolone acetonide, fluocortin butyl, fluocortolone, fluocortolone hexanoate, diflucortolone valerate, fluorometholone, fluperolone acetate, fluprednidene acetate, fluprednisolone, flurandrenolide, formocortol, halcinonide, halometasone, halopredone acetate, hydro-cortamate, hydrocortisone, hydrocortisone acetate, hydrocortisone butyrate, hydrocortisone phosphate, hydrocortisone 21-sodium succinate, hydrocortisone tebutate, mazipredone, medrysone, meprednisone, methylprednisolone, mometasone furoate, paramethasone, prednicarbate, prednisolone, prednisolone 21-diedryaminoacetate, prednisolone sodium phosphate, prednisolone sodium succinate, prednisolone sodium 21-*m*-sulfobenzoate, prednisolone sodium 21-stearoglycolate, prednisolone tebutate, prednisolone 21-trimethylacetate, prednisone, prednival, prednylidene, prednylidene 21-diethylaminoacetate, tixocortol, triamcinolone, triamcinolone acetonide, triamcinolone benetonide and triamcinolone hexacetonide.

Structurally related corticosteroids having similar analgesic and anti-inflammatory properties are also intended to be encompassed by this group.

In another specific embodiment, the present invention is directed to the use of an huE3 α polypeptide, or antagonist (including, but not limited to, anti-huE3 α selective binding agents) in combination (pretreatment, post-treatment, or concurrent treatment) with any of one or more slow-acting antirheumatic drugs (SAARDs) or disease modifying antirheumatic drugs (DMARDS), prodrug esters, or pharmaceutically acceptable salts thereof for the treatment of the diseases and disorders recited herein, including acute and chronic inflammation such as rheumatic diseases, graft versus host disease and multiple sclerosis. SAARDs or DMARDS, prodrug esters and pharmaceutically acceptable salts thereof comprise: allocupreide sodium, auranofin, aurothioglucose, aurothioglycanide, azathioprine, brequinar sodium, bucillamine, calcium 3-aurothio-2-propanol-1-sulfonate, chlorambucil, chloroquine, clobuzarit, cuproxoline, cyclo-phosphamide, cyclosporin, dapsone, 15-deoxyspergualin, diacerein, glucosamine, gold salts (*e.g.*, cycloquine gold salt, gold sodium thiomalate, gold sodium thiosulfate), hydroxychloroquine, hydroxychloroquine sulfate, hydroxyurea, kebuzone, levamisole, lobenzarit, melittin, 6-mercaptopurine, methotrexate, mizoribine, mycophenolate mofetil, myoral, nitrogen mustard, D-penicillamine, pyridinol imidazoles such as SKNF86002 and SB203580, rapamycin, thiols, thymopoietin and vincristine. Structurally related SAARDs or DMARDS having similar analgesic and anti-inflammatory properties are also intended to be encompassed by this group.

In another specific embodiment, the present invention is directed to the use of a huE3 α polypeptide, or antagonist (including, but not limited to, anti-huE3 α selective binding agents) in combination (pretreatment, post-treatment, or concurrent treatment) with any of one or more COX2 inhibitors, prodrug esters or pharmaceutically acceptable salts thereof for the treatment of the diseases and disorders recited herein, including acute and chronic inflammation. Examples of COX2 inhibitors, prodrug esters or pharmaceutically acceptable salts thereof include, for example, celecoxib. Structurally related COX2 inhibitors having similar analgesic and anti-inflammatory properties are also intended to be encompassed by this group.

In still another specific embodiment, the present invention is directed to the use of a huE3 α polypeptide, or antagonist (including, but not limited to, anti-huE3 α selective binding agents) in combination (pretreatment, post-treatment, or concurrent treatment) with any of one or more antimicrobials, prodrug esters or pharmaceutically

acceptable salts thereof for the treatment of the diseases and disorders recited herein, including cachexia, muscle wasting diseases and other catabolic disorders. Antimicrobials include, for example, the broad classes of penicillins, cephalosporins and other beta-lactams, aminoglycosides, azoles, quinolones, macrolides, rifamycins, tetracyclines, sulfonamides, lincosamides and polymyxins. The penicillins include, but are not limited to penicillin G, penicillin V, methicillin, nafcillin, oxacillin, cloxacillin, dicloxacillin, floxacillin, ampicillin, ampicillin/sulbactam, amoxicillin, amoxicillin/clavulanate, hetacillin, cyclacillin, bacampicillin, carbenicillin, carbenicillin indanyl, ticarcillin, ticarcillin/clavulanate, azlocillin, mezlocillin, peperacillin, and mecillinam. The cephalosporins and other beta-lactams include, but are not limited to cephalothin, cephapirin, cephalixin, cephradine, cefazolin, cefadroxil, cefaclor, cefamandole, cefotetan, cefoxitin, ceruroxime, cefonicid, ceforadine, cefixime, cefotaxime, moxalactam, ceftizoxime, ceftriaxone, cephooperazone, ceftazidime, imipenem and aztreonam. The aminoglycosides include, but are not limited to streptomycin, gentamicin, tobramycin, amikacin, netilmicin, kanamycin and neomycin. The azoles include, but are not limited to fluconazole. The quinolones include, but are not limited to nalidixic acid, norfloxacin, enoxacin, ciprofloxacin, ofloxacin, sparfloxacin and temafloxacin. The macrolides include, but are not limited to erythromycin, spiramycin and azithromycin. The rifamycins include, but are not limited to rifampin. The tetracyclines include, but are not limited to spicycline, chlortetracycline, clomocycline, demeclocycline, deoxycycline, guamecycline, lymecycline, meclocycline, methacycline, minocycline, oxytetracycline, penimepicycline, pipacycline, rolitetracycline, sancycline, senociclin and tetracycline. The sulfonamides include, but are not limited to sulfanilamide, sulfamethoxazole, sulfacetamide, sulfadiazine, sulfisoxazole and co-trimoxazole (trimethoprim/sulfamethoxazole). The lincosamides include, but are not limited to clindamycin and lincomycin. The polymyxins (polypeptides) include, but are not limited to polymyxin B and colistin.

Human E3 α Compositions and Administration

Therapeutic compositions are within the scope of the present invention. Such compositions may comprise a therapeutically effective amount of a huE3 α polypeptide, including a fragment, variant, derivative, or one or more selective binding agents which either inhibit or stimulate an activity of huE3 α in admixture with a

pharmaceutically acceptable agent such as a pharmaceutically acceptable formulation agent; wherein huE3 α refers to the polypeptide sequence of huE3 α I or huE3 α II.

Human E3 α pharmaceutical compositions typically include a therapeutically or prophylactically effective amount of huE3 α polypeptide, (an inhibitor of huE3 α action) nucleic acid molecule or selective binding agent in a mixture with one or more pharmaceutically and physiologically acceptable formulation agents selected for suitability with the mode of administration. Suitable formulation materials or pharmaceutically acceptable agents include, but are not limited to, antioxidants, preservatives, coloring, flavoring and diluting agents, emulsifying agents, suspending agents, solvents, fillers, bulking agents, buffers, delivery vehicles, diluents, excipients and/or pharmaceutical adjuvants. For example, a suitable vehicle or carrier may be water for injection, physiological saline solution, or artificial cerebrospinal fluid, possibly supplemented with other materials common in compositions for parenteral administration. Neutral buffered saline or saline mixed with serum albumin are further exemplary vehicles. The term "pharmaceutically acceptable carrier" or "physiologically acceptable carrier" as used herein refers to one or more formulation agents suitable for accomplishing or enhancing the delivery of the huE3 α polypeptide, nucleic acid molecule or selective binding agent as a pharmaceutical composition.

Acceptable formulation materials preferably are nontoxic to recipients and are preferably inert at the dosages and concentrations employed. The materials may include buffers such as phosphate, citrate, or other organic acids; antioxidants such as ascorbic acid; low molecular weight polypeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids such as glycine, glutamine, asparagine, arginine or lysine; monosaccharides, disaccharides, and other carbohydrates including glucose, mannose, or dextrans; chelating agents such as ethylenediamine tetraacetic acid (EDTA); sugar alcohols such as mannitol or sorbitol; salt-forming counterions such as sodium; and/or nonionic surfactants such as tween, pluronics, or polyethylene glycol (PEG).

Typically, a huE3 α molecule pharmaceutical composition will be administered in the form of a composition comprising a purified polypeptide, in conjunction with one or more physiologically acceptable agents. It will be appreciated that when used herein, the term "huE3 α molecule pharmaceutical composition" also encompasses compositions containing a nucleic acid molecule or selective binding agent of the present invention.

Neutral buffered saline or saline mixed with serum albumin are exemplary appropriate carriers. Other standard pharmaceutically acceptable agents such as diluents and excipients may be included as desired. For example, the huE3 α polypeptide product may be formulated as a lyophilizate using appropriate excipients such as sucrose. Other exemplary pharmaceutical compositions comprise Tris buffer of about pH 7.0-8.5, or acetate buffer of about pH 4.0-5.5, which may further include sorbitol or a suitable substitute therefor.

The primary vehicle or carrier in a pharmaceutical composition may be either aqueous or non-aqueous in nature. In addition, the composition may contain other formulation materials for modifying or maintaining the pH, osmolarity, viscosity, clarity, color, sterility, stability, rate of dissolution, or odor of the formulation. Similarly, the composition may contain additional formulation materials for modifying or maintaining the rate of release of huE3 α polypeptide, nucleic acid molecule or selective binding agent, or for promoting the absorption or penetration of huE3 α such molecules.

The huE3 α molecule pharmaceutical compositions can be administered parenterally. Alternatively, the compositions may be administered through the digestive tract, such as orally, or by inhalation. When parenterally administered, the therapeutic compositions for use in this invention may be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such pharmaceutically acceptable compositions, with due regard to pH, isotonicity, stability and the like, is within the skill of the art.

A particularly suitable vehicle for parenteral injection is sterile distilled water in which a huE3 α polypeptide is formulated as a sterile, isotonic solution, properly preserved. Yet another preparation can involve the formulation of the desired molecule with an agent, such as injectable microspheres, bio-erodible particles or beads, or liposomes, that provides for the controlled or sustained release of the product which may then be delivered as a depot injection. Other suitable means for the introduction of the desired molecule include implantable drug delivery devices.

The pharmaceutical compositions of the present invention may include other components, for example parenterally acceptable preservatives, tonicity agents, cosolvents, wetting agents, complexing agents, buffering agents, antimicrobials, antioxidants and surfactants, as are well known in the art. For example, suitable tonicity enhancing agents include alkali metal halides (preferably sodium or potassium chloride),

mannitol, sorbitol, and the like. Suitable preservatives include, but are not limited to, benzalkonium chloride, thimerosal, phenethyl alcohol, methylparaben, propylparaben, chlorhexidine, sorbic acid, and the like. Hydrogen peroxide may also be used as preservative. Suitable cosolvents are for example glycerin, propylene glycol and polyethylene glycol. Suitable complexing agents are for example caffeine, polyvinylpyrrolidone, beta-cyclodextrin or hydroxypropyl-beta-cyclodextrin. Suitable surfactants or wetting agents include sorbitan esters, polysorbates such as polysorbate 80, tromethamine, lecithin, cholesterol, tyloxapal, and the like. The buffers can be conventional buffers such as borate, citrate, phosphate, bicarbonate, or Tris-HCl.

The formulation components are present in concentrations that are acceptable to the site of administration. For example, buffers are used to maintain the composition at physiological pH or at slightly lower pH, typically within a pH range of from about 5 to about 8.

In one embodiment of the present invention, huE3 α polypeptide compositions may be prepared for storage by mixing the selected composition having the desired degree of purity with optional physiologically acceptable carriers, excipients, or stabilizers (*Remington's pharmaceutical sciences*, 18th edition, A.R. Gennaro, ed., Mack Publishing Company (1990)) in the form of a lyophilized cake or an aqueous solution.

The optimal pharmaceutical formulation will be determined by one skilled in the art depending upon, for example, the intended route of administration, delivery format, and desired dosage. See for example, *Remington's Pharmaceutical Sciences*, pp. 1435-1712. Such compositions may influence the physical state, stability, rate of *in vivo* release, and rate of *in vivo* clearance of the present huE3 α polypeptides.

An effective amount of a huE3 α polypeptide composition to be employed therapeutically will depend, for example, upon the therapeutic objectives such as the indication for which the huE3 α polypeptide is being used, the route of administration, and the condition of the patient. Accordingly, the clinician may titer the dosage and modify the route of administration to obtain the optimal therapeutic effect. A typical dosage may range from about 0.1 μ g/kg to up to about 100 mg/kg or more, depending on the factors mentioned above. In other embodiments, the dosage may range from 1 μ g/kg up to about 100 mg/kg; or 5 μ g/kg up to about 100 mg/kg; or 0.1 μ g/kg up to about 100 mg/kg; or 1 μ g/kg up to about 100 mg/kg.

Typically, a clinician will administer the composition until a dosage is reached that achieves the desired effect. The composition may therefore be administered as a single dose, or as two or more doses (which may or may not contain the same amount of the desired molecule) over time, or as a continuous infusion via implantation device or catheter.

One skilled in the art will appreciate that the appropriate dosage levels for treatment will thus vary depending, in part, upon the molecule delivered, the therapeutic context, type of disorder under treatment, the age, and general health of the recipient.

The huE3 α molecule pharmaceutical composition to be used for *in vivo* administration typically must be sterile. This may be accomplished by filtration through sterile filtration membranes. Where the composition is lyophilized, sterilization using these methods may be conducted either prior to, or following, lyophilization and reconstitution. The composition for parenteral administration may be stored in lyophilized form or in solution. In addition, parenteral compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

Once the pharmaceutical composition has been formulated, it may be stored in sterile vials as a solution, suspension, gel, emulsion, solid, or a dehydrated or lyophilized powder. Such formulations may be stored either in a ready-to-use form or in a form (*e.g.*, lyophilized) requiring reconstitution prior to administration.

In a specific embodiment, the present invention is directed to kits for producing a single-dose administration unit. The kits may each contain both a first container having a dried protein and a second container having an aqueous formulation. Also included within the scope of this invention are kits containing single and multi-chambered pre-filled syringes (*e.g.*, liquid syringes and lyosyringes).

An effective amount of a h2520-109 pharmaceutical composition to be employed therapeutically will depend, for example, upon the therapeutic context and objectives. One skilled in the art will appreciate that the appropriate dosage levels for treatment will thus vary depending, in part, upon the molecule delivered, the indication for which the h2520-109 molecule is being used, the route of administration, and the size (body weight, body surface or organ size) and condition (the age and general health) of the patient. Accordingly, the clinician may titer the dosage and modify the route of administration to obtain the optimal therapeutic effect. A typical dosage may range

from about 0.1 mg/kg to up to about 100 mg/kg or more, depending on the factors mentioned above. In other embodiments, the dosage may range from 0.1 mg/kg up to about 100 mg/kg; or 1 mg/kg up to about 100 mg/kg; or 5 mg/kg up to about 100 mg/kg.

The frequency of dosing will depend upon the pharmacokinetic parameters of the h2520-109 molecule in the formulation used. Typically, a clinician will administer the composition until a dosage is reached that achieves the desired effect. The composition may therefore be administered as a single dose, or as two or more doses (which may or may not contain the same amount of the desired molecule) over time, or as a continuous infusion via implantation device or catheter.

Pharmaceutical compositions such as (1) slow-release formulations, (2) inhalant mists, or (3) orally active formulations are also envisioned. The huE3 α molecule pharmaceutical composition generally is formulated for parenteral administration. Such parenterally administered therapeutic compositions are typically in the form of a pyrogen-free, parenterally acceptable aqueous solution comprising the desired huE3 α molecule in a pharmaceutically acceptable vehicle. The huE3 α molecule pharmaceutical compositions also may include particulate preparations of polymeric compounds such as polylactic acid, polyglycolic acid, etc. or the introduction of the molecule into liposomes. Hyaluronic acid may also be used, and this may have the effect of promoting sustained duration in the circulation.

In one embodiment, a pharmaceutical composition may be formulated for inhalation. For example, huE3 α polypeptide may be formulated as a dry powder for inhalation. Human E3 α polypeptide or nucleic acid molecule inhalation solutions may also be formulated in a liquefied propellant for aerosol delivery, with or without a liquified propellant. In yet another embodiment, solutions may be nebulized. Pulmonary administration is further described in PCT WO94/20069, which describes pulmonary delivery of chemically modified proteins.

It is also contemplated that certain formulations may be administered orally. In one embodiment of the present invention, huE3 α polypeptides which are administered in this fashion can be formulated with or without those carriers customarily used in the compounding of solid dosage forms such as tablets and capsules. For example, a capsule may be designed to release the active portion of the formulation at the point in the gastrointestinal tract when bioavailability is maximized and pre-systemic degradation is minimized. Additional agents can be included to facilitate absorption of the huE3 α polypeptide. Diluents, flavorings, low melting point waxes, vegetable oils,

lubricants, suspending agents, tablet disintegrating agents, and binders may also be employed.

Another pharmaceutical composition may involve an effective quantity of huE3 α polypeptides in a mixture with non-toxic excipients which are suitable for the manufacture of tablets. By dissolving the tablets in sterile water, or other appropriate vehicle, solutions can be prepared in unit dose form. Suitable excipients include, but are not limited to, inert diluents, such as calcium carbonate, sodium carbonate or bicarbonate, lactose, or calcium phosphate; or binding agents, such as starch, gelatin, or acacia; or lubricating agents such as magnesium stearate, stearic acid, or talc.

Additional huE3 α molecule formulations will be evident to those skilled in the art, including formulations involving huE3 α molecules in combination with one or more other therapeutic agents. Techniques for formulating a variety of other sustained- or controlled-delivery means, such as liposome carriers, bio-erodible microparticles or porous beads and depot injections, are also known to those skilled in the art. See for example, PCT/US93/00829 which describes controlled release of porous polymeric microparticles for the delivery of pharmaceutical compositions.

Additional examples of sustained-release preparations include semipermeable polymer matrices in the form of shaped articles, *e.g.* films, or microcapsules. Sustained release matrices may include polyesters, hydrogels, polylactides (U.S. 3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma ethyl-L-glutamate (Sidman *et al.*, *Biopolymers*, 22: 547-556, 1983), poly (2-hydroxyethyl-methacrylate) (Langer *et al.*, *J. Biomed. Mater. Res.*, 15: 167-27, 1981; and Langer, *Chem. Tech.*, 12: 98-105, 1982), ethylene vinyl acetate (Langer *et al.*, *supra*) or poly-D(-)-3-hydroxybutyric acid (EP 133,988). Sustained-release compositions also may include liposomes, which can be prepared by any of several methods known in the art. (See *e.g.*, Eppstein *et al.*, *Proc. Natl. Acad. Sci. USA*, 82: 3688-3692, 1985; EP 36,676; EP 88,046; EP 143,949.)

Regardless of the manner of administration, the specific dose may be calculated according to body weight, body surface area or organ size. Further refinement of the appropriate dosage is routinely made by those of ordinary skill in the art and is within the ambit of tasks routinely performed by them. Appropriate dosages may be ascertained through use of appropriate dose-response data.

The route of administration of the pharmaceutical composition is in accord with known methods, *e.g.* oral, inhalation, injection or infusion by intravenous,

intraperitoneal, intracerebral (intra-parenchymal), intracerebroventricular, intramuscular, intra-ocular, intraarterial, intraportal, or intralesional routes, or by sustained release systems or implantation device. Where desired, the compositions may be administered continuously by infusion, by bolus injection devices or by implantation device.

5 Alternatively or additionally, the composition may be administered locally via implantation into the affected area of a membrane, sponge, or other appropriate material on to which the desired molecule has been absorbed or encapsulated. Where an implantation device is used, the device may be implanted into any suitable tissue or organ, and delivery of the desired molecule may be directly
10 through the device via diffusion, time-released bolus, or via continuous administration, or via catheter using continuous infusion.

It will further be appreciated that the huE3 α polypeptides, including fragments, variants, and derivatives, may be employed alone, together, or in combination with other polypeptides and pharmaceutical compositions. For example, the huE3 α
15 polypeptides may be used in combination with cytokines, growth factors, antibiotics, anti-inflammatories, and/or chemotherapeutic agents as is appropriate for the indication being treated.

In some cases, it may be desirable to use huE3 α pharmaceutical compositions in an *ex vivo* manner. In such instances, cells, tissues, or organs that have
20 been removed from the patient are exposed to huE3 α pharmaceutical compositions after which the cells, tissues and/or organs are subsequently implanted back into the patient.

In other cases, a huE3 α polypeptide can be delivered by implanting certain cells that have been genetically engineered, using methods such as those described herein, to express and secrete the polypeptides. Such cells may be animal or
25 human cells, and may be autologous, heterologous, or xenogeneic. Optionally, the cells may be immortalized. However, in order to decrease the chance of an immunological response, the cells may be encapsulated to avoid infiltration of surrounding tissues. The encapsulation materials are typically biocompatible, semi-permeable polymeric enclosures or membranes that allow the release of the protein product(s) but prevent the
30 destruction of the cells by the patient's immune system or by other detrimental factors from the surrounding tissues.

Additional embodiments of the present invention relate to cells and methods (*e.g.*, homologous recombination and /or other recombinant production methods) for both the *in vitro* production of therapeutic polypeptides by means of

homologous recombination and for the production and delivery of therapeutic polypeptides by gene therapy or cell therapy.

It is further envisioned that huE3 α polypeptides can be produced by homologous recombination, or with recombinant production methods utilizing control elements introduced into cells already containing DNA encoding huE3 α polypeptides. For example, homologous recombination methods may be used to modify a cell that contains a normally transcriptionally silent huE3 α gene, or an under expressed gene, and thereby produce a cell which expresses therapeutically efficacious amounts of huE3 α polypeptides. Homologous recombination is a technique originally developed for targeting genes to induce or correct mutations in transcriptionally active genes. Kucherlapati, *Prog. in Nucl. Acid Res. & Mol. Biol.*, 36:301, 1989. The basic technique was developed as a method for introducing specific mutations into specific regions of the mammalian genome (Thomas *et al.*, *Cell*, 44: 419-428, 1986; Thomas and Capecchi, *Cell*, 51:503-512, 1987; Doetschman *et al.*, *Proc. Natl. Acad. Sci.*, 85: 8583-8587, 1988) or to correct specific mutations within defective genes (Doetschman *et al.*, *Nature*, 330: 576-578, 1987). Exemplary homologous recombination techniques are described in U.S. Patent No. 5,272,071 (EP 9193051, EP Publication No. 505500; PCT/US90/07642, International Publication No. WO 91/09955).

Through homologous recombination, the DNA sequence to be inserted into the genome can be directed to a specific region of the gene of interest by attaching it to targeting DNA. The targeting DNA is a nucleotide sequence that is complementary (homologous) to a region of the genomic DNA. Small pieces of targeting DNA that are complementary to a specific region of the genome are put in contact with the parental strand during the DNA replication process. It is a general property of DNA that has been inserted into a cell to hybridize, and therefore, recombine with other pieces of endogenous DNA through shared homologous regions. If this complementary strand is attached to an oligonucleotide that contains a mutation or a different sequence or an additional nucleotide, it too is incorporated into the newly synthesized strand as a result of the recombination. As a result of the proofreading function, it is possible for the new sequence of DNA to serve as the template. Thus, the transferred DNA is incorporated into the genome.

Attached to these pieces of targeting DNA are regions of DNA which may interact with or control the expression of a huE3 α polypeptide, *e.g.*, flanking sequences. For example, a promoter/enhancer element, a suppressor, or an exogenous

transcription modulatory element is inserted in the genome of the intended host cell in proximity and orientation sufficient to influence the transcription of DNA encoding the desired huE3 α polypeptide. The control element controls a portion of the DNA present in the host cell genome. Thus, the expression of huE3 α polypeptide may be achieved not by transfection of DNA that encodes the huE3 α gene itself, but rather by the use of targeting DNA (containing regions of homology with the endogenous gene of interest) coupled with DNA regulatory segments that provide the endogenous gene sequence with recognizable signals for transcription of a huE3 α polypeptide.

In an exemplary method, the expression of a desired targeted gene in a cell (*i.e.*, a desired endogenous cellular gene) is altered by the introduction, by homologous recombination into the cellular genome at a preselected site, by the introduction of DNA which includes at least a regulatory sequence, an exon and a splice donor site. These components are introduced into the chromosomal (genomic) DNA in such a manner that this, in effect, results in the production of a new transcription unit (in which the regulatory sequence, the exon and the splice donor site present in the DNA construct are operatively linked to the endogenous gene). As a result of the introduction of these components into the chromosomal DNA, the expression of the desired endogenous gene is altered.

Altered gene expression, as described herein, encompasses activating (or causing to be expressed) a gene which is normally silent (unexpressed) in the cell as obtained, as well as increasing the expression of a gene which is not expressed at physiologically significant levels in the cell as obtained. The embodiments further encompass changing the pattern of regulation or induction such that it is different from the pattern of regulation or induction that occurs in the cell as obtained, and reducing (including eliminating) the expression of a gene which is expressed in the cell as obtained.

One method by which homologous recombination can be used to increase, or cause, huE3 α polypeptide production from a cell's endogenous huE3 α gene involves first using homologous recombination to place a recombination sequence from a site-specific recombination system (*e.g.*, Cre/loxP, FLP/FRT) (Sauer, *Current Opinion In Biotechnology*, 5: 521-527, 1994; Sauer, *Methods In Enzymology*, 225: 890-900, 1993) upstream (that is, 5' to) of the cell's endogenous genomic huE3 α coding region. A plasmid containing a recombination site homologous to the site that was placed just upstream of the genomic huE3 α coding region is introduced into the modified cell line

along with the appropriate recombinase enzyme. This recombinase causes the plasmid to integrate, via the plasmid's recombination site, into the recombination site located just upstream of the genomic huE3 α coding region in the cell line (Baubonis and Sauer, *Nucleic Acids Res.*, 21: 2025-2029, 1993; O'Gorman *et al.*, *Science*, 251: 1351-1355, 1991). Any flanking sequences known to increase transcription (*e.g.*, enhancer/promoter, intron, translational enhancer), if properly positioned in this plasmid, would integrate in such a manner as to create a new or modified transcriptional unit resulting in *de novo* or increased huE3 α polypeptide production from the cell's endogenous huE3 α gene.

A further method to use the cell line in which the site specific recombination sequence had been placed just upstream of the cell's endogenous genomic huE3 α coding region is to use homologous recombination to introduce a second recombination site elsewhere in the cell line's genome. The appropriate recombinase enzyme is then introduced into the two-recombination-site cell line, causing a recombination event (deletion, inversion, translocation) (Sauer, *Current Opinion In Biotechnology*, 5: 521-527, 1994; Sauer, *Methods In Enzymology*, 225: 890-900, 1993) that would create a new or modified transcriptional unit resulting in *de novo* or increased huE3 α polypeptide production from the cell's endogenous huE3 α gene.

An additional approach for increasing, or causing, the expression of huE3 α polypeptide from a cell's endogenous huE3 α gene involves increasing, or causing, the expression of a gene or genes (*e.g.*, transcription factors) and/or decreasing the expression of a gene or genes (*e.g.*, transcriptional repressors) in a manner which results in *de novo* or increased huE3 α polypeptide production from the cell's endogenous huE3 α gene. This method includes the introduction of a non-naturally occurring polypeptide (*e.g.*, a polypeptide comprising a site specific DNA binding domain fused to a transcriptional factor domain) into the cell such that *de novo* or increased huE3 α polypeptide production from the cell's endogenous huE3 α gene results.

The present invention further relates to DNA constructs useful in the method of altering expression of a target gene. In certain embodiments, the exemplary DNA constructs comprise: (a) one or more targeting sequences; (b) a regulatory sequence; (c) an exon; and (d) an unpaired splice-donor site. The targeting sequence in the DNA construct directs the integration of elements (a)-(d) into a target gene in a cell such that the elements (b)-(d) are operatively linked to sequences of the endogenous target gene. In another embodiment, the DNA constructs comprise: (a) one or more

targeting sequences, (b) a regulatory sequence, (c) an exon, (d) a splice-donor site, (e) an intron, and (f) a splice-acceptor site, wherein the targeting sequence directs the integration of elements (a)-(f) such that the elements of (b)-(f) are operatively linked to the endogenous gene. The targeting sequence is homologous to the preselected site in the cellular chromosomal DNA with which homologous recombination is to occur. In the construct, the exon is generally 3' of the regulatory sequence and the splice-donor site is 3' of the exon.

If the sequence of a particular gene is known, such as the nucleic acid sequence encoding a huE3 α polypeptide presented herein, a piece of DNA that is complementary to a selected region of the gene can be synthesized or otherwise obtained, such as by appropriate restriction of the native DNA at specific recognition sites bounding the region of interest. This piece serves as a targeting sequence upon insertion into the cell and will hybridize to its homologous region within the genome. If this hybridization occurs during DNA replication, this piece of DNA, and any additional sequence attached thereto, will act as an Okazaki fragment and will be incorporated into the newly synthesized daughter strand of DNA. The present invention, therefore, includes nucleotides encoding a huE3 α polypeptide, which nucleotides may be used as targeting sequences.

Human E3 α polypeptide cell therapy, *e.g.*, the implantation of cells producing huE3 α polypeptides, is also contemplated. This embodiment involves implanting cells capable of synthesizing and secreting a biologically active form of huE3 α polypeptide. Such huE3 α polypeptide-producing cells can be cells that are natural producers of huE3 α polypeptides or may be recombinant cells whose ability to produce huE3 α polypeptides has been augmented by transformation with a gene encoding the desired huE3 α polypeptide or with a gene augmenting the expression of huE3 α polypeptide. Such a modification may be accomplished by means of a vector suitable for delivering the gene as well as promoting its expression and secretion. In order to minimize a potential immunological reaction in patients being administered a huE3 α polypeptide, as may occur with the administration of a polypeptide of a foreign species, it is preferred that the natural cells producing huE3 α polypeptide be of human origin and produce huE3 α polypeptide. Likewise, it is preferred that the recombinant cells producing huE3 α polypeptide be transformed with an expression vector containing a gene encoding a human huE3 α polypeptide.

Implanted cells may be encapsulated to avoid the infiltration of surrounding tissue. Human or non-human animal cells may be implanted in patients in biocompatible, semipermeable polymeric enclosures or membranes that allow the release of huE3 α polypeptide, but that prevent the destruction of the cells by the patient's immune system or by other detrimental factors from the surrounding tissue. Alternatively, the patient's own cells, transformed to produce huE3 α polypeptides *ex vivo*, may be implanted directly into the patient without such encapsulation.

Techniques for the encapsulation of living cells are known in the art, and the preparation of the encapsulated cells and their implantation in patients may be routinely accomplished. For example, Baetge *et al.* (WO95/05452; PCT/US94/09299) describe membrane capsules containing genetically engineered cells for the effective delivery of biologically active molecules. The capsules are biocompatible and are easily retrievable. The capsules encapsulate cells transfected with recombinant DNA molecules comprising DNA sequences coding for biologically active molecules operatively linked to promoters that are not subject to down regulation *in vivo* upon implantation into a mammalian host. The devices provide for the delivery of the molecules from living cells to specific sites within a recipient. In addition, *see* U.S. Patent Nos. 4,892,538, 5,011,472, and 5,106,627. A system for encapsulating living cells is described in PCT Application WO91/10425 of Aebischer *et al.* *See also*, PCT Application WO91/10470 of Aebischer *et al.*, Winn *et al.*, *Exper. Neurol.*, 113: 322-329, 1991, Aebischer *et al.*, *Exper. Neurol.*, 111: 269-275, 1991; and Tresco *et al.*, *ASAIO*, 38: 17-23, 1992.

In vivo and *in vitro* gene therapy delivery of huE3 α polypeptides is also envisioned. *In vivo* gene therapy may be accomplished by introducing the gene encoding huE3 α polypeptide into cells via local injection of a huE3 α nucleic acid molecule or by other appropriate viral or non-viral delivery vectors (Hefti, *Neurobiology*, 25: 1418-1435, 1994). For example, a nucleic acid molecule encoding a huE3 α polypeptide may be contained in an adeno-associated virus vector for delivery to the targeted cells (*e.g.*, Johnson, International Publication No. WO95/34670; International Application No. PCT/US95/07178). The recombinant adeno-associated virus (AAV) genome typically contains AAV inverted terminal repeats flanking a DNA sequence encoding a huE3 α polypeptide operably linked to functional promoter and polyadenylation sequences.

Alternative suitable viral vectors include, but are not limited to, retrovirus, adenovirus, herpes simplex virus, lentivirus, hepatitis virus, parvovirus, papovavirus, poxvirus, alphavirus, coronavirus, rhabdovirus, paramyxovirus, and papilloma virus vectors. U.S. Patent No. 5,672,344 describes an *in vivo* viral-mediated gene transfer system involving a recombinant neurotrophic HSV-1 vector. U.S. Patent No. 5,399,346 provides examples of a process for providing a patient with a therapeutic protein by the delivery of human cells which have been treated *in vitro* to insert a DNA segment encoding a therapeutic protein. Additional methods and materials for the practice of gene therapy techniques are described in U.S. Patent No. 5,631,236 involving adenoviral vectors; U.S. Patent No. 5,672,510 involving retroviral vectors; and U.S. 5,635,399 involving retroviral vectors expressing cytokines.

Nonviral delivery methods include, but are not limited to, liposome-mediated transfer, naked DNA delivery (direct injection), receptor-mediated transfer (ligand-DNA complex), electroporation, calcium phosphate precipitation, and microparticle bombardment (*e.g.*, gene gun). Gene therapy materials and methods may also include the use of inducible promoters, tissue-specific enhancer-promoters, DNA sequences designed for site-specific integration, DNA sequences capable of providing a selective advantage over the parent cell, labels to identify transformed cells, negative selection systems and expression control systems (safety measures), cell-specific binding agents (for cell targeting), cell-specific internalization factors, and transcription factors to enhance expression by a vector as well as methods of vector manufacture. Such additional methods and materials for the practice of gene therapy techniques are described in U.S. Patent No. 4,970,154 involving electroporation techniques; WO96/40958 involving nuclear ligands; U.S. Patent No. 5,679,559 describing a lipoprotein-containing system for gene delivery; U.S. Patent No. 5,676,954 involving liposome carriers; U.S. Patent No. 5,593,875 concerning methods for calcium phosphate transfection; and U.S. Patent No. 4,945,050 wherein biologically active particles are propelled at cells at a speed whereby the particles penetrate the surface of the cells and become incorporated into the interior of the cells.

In yet other embodiments, regulatory elements can be included for the controlled expression of the huE3 α gene in the target cell. Such elements are turned on in response to an appropriate effector. In this way, a therapeutic polypeptide can be expressed when desired. One conventional control means involves the use of small molecule dimerizers or rapalogs (as described in WO9641865 (PCT/US96/099486);

WO9731898 (PCT/US97/03137) and WO9731899 (PCT/US95/03157)) used to dimerize chimeric proteins which contain a small molecule-binding domain and a domain capable of initiating biological process, such as a DNA-binding protein or transcriptional activation protein. The dimerization of the proteins can be used to initiate transcription of the huE3 α gene.

Other suitable control means or gene switches include, but are not limited to, the following systems. Mifepristone (RU486) is used as a progesterone antagonist. The binding of a modified progesterone receptor ligand-binding domain to the progesterone antagonist activates transcription by forming a dimer of two transcription factors which then pass into the nucleus to bind DNA. The ligand binding domain is modified to eliminate the ability of the receptor to bind to the natural ligand. The modified steroid hormone receptor system is further described in U.S. 5,364,791; WO9640911, and WO9710337.

Yet another control system uses ecdysone (a fruit fly steroid hormone) which binds to and activates an ecdysone receptor (cytoplasmic receptor). The receptor then translocates to the nucleus to bind a specific DNA response element (promoter from ecdysone-responsive gene). The ecdysone receptor includes a transactivation domain/DNA-binding domain/ligand-binding domain to initiate transcription. The ecdysone system is further described in U.S. 5,514,578; WO9738117; WO9637609; and WO9303162.

Another control means uses a positive tetracycline-controllable transactivator. This system involves a mutated tet repressor protein DNA-binding domain (mutated tet R-4 amino acid changes which resulted in a reverse tetracycline-regulated transactivator protein, *i.e.*, it binds to a tet operator in the presence of tetracycline) linked to a polypeptide which activates transcription. Such systems are described in U.S. Patent Nos. 5,464,758; 5,650,298 and 5,654,168.

Additional expression control systems and nucleic acid constructs are described in U.S. Patent Nos. 5,741,679 and 5,834,186 to Innovir Laboratories Inc.

One example of a gene therapy technique is to use the huE3 α gene (either genomic DNA, cDNA, and/or synthetic DNA encoding a huE3 α polypeptide which may be operably linked to a constitutive or inducible promoter to form a "gene therapy DNA construct". The promoter may be homologous or heterologous to the endogenous huE3 α gene, provided that it is active in the cell or tissue type into which the construct will be inserted. Other components of the gene therapy DNA construct may optionally include,

DNA molecules designed for site-specific integration (*e.g.*, endogenous sequences useful for homologous recombination), tissue-specific promoter, enhancer(s) or silencer(s), DNA molecules capable of providing a selective advantage over the parent cell, DNA molecules useful as labels to identify transformed cells, negative selection systems, cell specific binding agents (as, for example, for cell targeting), cell-specific internalization factors, and transcription factors to enhance expression by a vector as well as factors to enable vector manufacture.

This gene therapy DNA construct can then be introduced into cells (either *ex vivo* or *in vivo*). One means for introducing the gene therapy DNA construct is by means of viral vectors as described herein. Certain vectors, such as retroviral vectors, will deliver the gene therapy DNA construct to the chromosomal DNA of the cells, and the gene therapy DNA construct can integrate into the chromosomal DNA. Other vectors will function as episomes, and the gene therapy DNA construct will remain in the cytoplasm.

Another means to increase endogenous huE3 α polypeptide expression in a cell via gene therapy is to insert one or more enhancer elements into the huE3 α polypeptide promoter, where the enhancer element(s) can serve to increase transcriptional activity of the huE3 α gene. The enhancer element(s) used will be selected based on the tissue in which one desires to activate the gene(s); enhancer elements known to confer promoter activation in that tissue will be selected. For example, if a gene encoding a huE3 α polypeptide is to be “turned on” in T-cells, the *lck* promoter enhancer element may be used. Here, the functional portion of the transcriptional element to be added may be inserted into a fragment of DNA containing the huE3 α polypeptide promoter (and optionally, inserted into a vector and/or 5’ and/or 3’ flanking sequence(s), etc.) using standard cloning techniques. This construct, known as a “homologous recombination construct”, can then be introduced into the desired cells either *ex vivo* or *in vivo*.

Gene therapy can be used to decrease huE3 α polypeptide expression by modifying the nucleotide sequence of the endogenous promoter(s). Such modification is typically accomplished via homologous recombination methods. For example, a DNA molecule containing all or a portion of the promoter of the huE3 α gene(s) selected for inactivation can be engineered to remove and/or replace pieces of the promoter that regulate transcription. For example the TATA box and/or the binding site of a transcriptional activator of the promoter may be deleted using standard molecular

biology techniques; such deletion can inhibit promoter activity thereby repressing the transcription of the corresponding huE3 α gene. The deletion of the TATA box or the transcription activator binding site in the promoter may be accomplished by generating a DNA construct comprising all or the relevant portion of the huE3 α polypeptide promoter(s) (from the same or a related species as the huE3 α gene(s) to be regulated) in which one or more of the TATA box and/or transcriptional activator binding site nucleotides are mutated via substitution, deletion and/or insertion of one or more nucleotides. As a result, the TATA box and/or activator binding site has decreased activity or is rendered completely inactive. This construct, which also will typically contain at least about 500 bases of DNA that correspond to the native (endogenous) 5' and 3' DNA sequences adjacent to the promoter segment that has been modified, may be introduced into the appropriate cells (either *ex vivo* or *in vivo*) either directly or via a viral vector as described herein. Typically, the integration of the construct into the genomic DNA of the cells will be via homologous recombination, where the 5' and 3' DNA sequences in the promoter construct can serve to help integrate the modified promoter region via hybridization to the endogenous chromosomal DNA.

Other gene therapy methods may also be employed where it is desirable to inhibit the activity of one or more huE3 α polypeptides. For example, antisense DNA or RNA molecules, which have a sequence that is complementary to at least a portion of the selected huE3 α gene(s) can be introduced into the cell. Typically, each such antisense molecule will be complementary to the start site (5' end) of each selected huE3 α gene. When the antisense molecule then hybridizes to the corresponding huE3 α mRNA, translation of this mRNA is prevented or reduced. It will also be appreciated by those skilled in the art that antisense and ribozyme molecules may also be administered directly.

Alternatively, gene therapy may be employed to create a dominant-negative inhibitor of one or more huE3 α polypeptides. In this situation, the DNA encoding a mutant full length or truncated polypeptide of each selected huE3 α polypeptide can be prepared and introduced into the cells of a patient using either viral or non-viral methods as described herein. Each such mutant is typically designed to compete with endogenous polypeptide in its biological role.

Additional Uses of huE3 α Nucleic Acids and Polypeptides

Nucleic acid molecules of the present invention may be used to map the locations of the huE3 α gene and related genes on chromosomes. Mapping may be done by techniques known in the art, such as PCR amplification and *in situ* hybridization.

5 The nucleic acid molecules are also used as antisense inhibitors of huE3 α polypeptide expression. Such inhibition may be effected by nucleic acid molecules which are complementary to and hybridize to expression control sequences (triple helix formation) or to huE3 α mRNA. Antisense probes may be designed by available techniques using the sequence of huE3 α nucleic acid molecules disclosed herein.
10 Antisense inhibitors provide information relating to the decrease or absence of a huE3 α polypeptide in a cell or organism.

 Hybridization probes may be prepared using the huE3 α nucleic acid sequences provided herein to screen cDNA, genomic or synthetic DNA libraries for related sequences. Regions of the DNA and/or amino acid sequence of huE3 α
15 polypeptide that exhibit significant identity to known sequences are readily determined using sequence alignment algorithms as described herein and those regions may be used to design probes for screening.

 Human E3 α nucleic acid molecules, as well as fragments, variants, and/or derivatives that do not themselves encode biologically active polypeptides, may be
20 useful as hybridization probes in diagnostic assays to test, either qualitatively or quantitatively, for the presence of huE3 α DNA or corresponding RNA in mammalian tissue or bodily fluid samples.

 Human E3 α polypeptide fragments, variants, and/or derivatives, whether biologically active or not, are also useful for preparing antibodies that bind to a huE3 α
25 polypeptide. The antibodies may be used for *in vitro* diagnostic purposes, including, but not limited to, use in labeled form to detect the presence of huE3 α polypeptide in a body fluid or cell sample.

 The full length cDNAs encoding huE3 α I was subcloned into pCR 2.1 vector (Invitrogen, Cat.# K2030-40). The full length cDNA encoding huE3 α II was
30 subcloned into pcDNA 3.1/His A vector (Invitrogen Cat.#V38-20). The full length cDNA encoding muE3 α II was subcloned into pCR 2.1 vector (Invitrogen). The above plasmids were deposited on March 15, 2000 to the American Type Culture Collection, 10801 University Boulevard, Manassas, VA 20110-2209 . The plasmid containing

huE3 α I is designated PTA-1489, the plasmid containing huE3 α II is designated PTA-1490 and the plasmid containing muE3 α II is designated PTA-1488.

The following examples are intended for illustration purposes only, and should not be construed as limiting the scope of the invention in any way.

EXAMPLE 1

Cloning of cDNA Encoding Human E3 α I

Materials and methods for cDNA cloning and analysis are described in Sambrook *et al.*, *supra*. which is incorporated herein by reference.

BLAST analysis of the Genebank dbEST database with the full length murine E3 α ubiquitin ligase nucleotide sequence (muE3I; Genebank Accession No.: AF061555; SEQ ID NO: 15), revealed 4 human EST sequences (Genebank accession numbers AI187306, AI92195, AI87306, and AI400279) which potentially encode different regions of a novel human E3 α ubiquitin ligase ortholog (huE3 α I) gene.

Based on these EST sequences, two sets of PCR primers (#2282-91/2282-93 and #2282-94/2282-97) were designed. These sequences are set out below in Table III.

Table III

Primer	Sequence	SEQ ID NO:
2282-91	CTC CTC GAG TCT GCG TCA AAC	7
2385-35	TCT GCA TAT GTT CAG CCT TGC TA	8
2282-94	GTA TGA ACT TGC CGA GGC TTT TA	9
2294-37	CAA TAC TTT CCC AGC CCT CAG AA	10

The primer sets #2282-91/2282-93 (SEQ ID NOS: 7 and 8) and #2282-94/2294-37 (SEQ ID NOS: 9 and 10) were used to generate two PCR products which span the whole huE3 α I gene including the 5' and 3' untranslated regions. Polymerase chain reactions (PCR) were performed using a Perkin-Elmer 9600 thermocycler. In general, 50 μ l PCR reactions contained 24 μ l of H₂O, 5 μ l of 10x cDNA PCR Reaction Buffer (Clontech), 2 μ l of 10 mM dNTP mix (dATP, dCTP, dGTP, dTTP), 1 μ l of Primer 2282-91 or 2282-94 (20 μ l), 1 μ l of Primer 2285-35 or 2294-37 (20 μ l), 2 μ l of 50x Advantage 2 Polymerase Mix (Clontech) and 15 μ l of Marathon Ready cDNA from a human heart library (Clontech cat.# 7404-1) or a human muscle library (Clontech cat. # 7413-1). The reaction mixture was incubated at 94°C for 30 seconds, followed by 40 cycles of 94°C for 30 seconds, 60°C for 30 seconds, and 72°C for 5 minutes.

The PCR products were electrophoresed on a 1% agarose gel as described by Sambrook *et al.*, *supra*. The appropriate sized bands (14 kB and 3 kB) were excised from the agarose gel and purified with the QIAquick Gel Extraction kit (Qiagen, cat# 28704). The two purified DNA fragments were subcloned into pCR2.1 vectors and transformed into *E. coli* (Strain INV α F) utilizing the Invitrogen Original TA Cloning kit (cat.# K2000-40).

After subcloning, DNA plasmids were purified with the QIAprep Spin Miniprep kit (cat# 27104). The sequence of the PCR products were verified by automated sequencing with the Prism 377 Sequencer and the Big Dye Terminator Ready Reaction mix with AmpliTaq DNA polymerase (Perkin Elmer Applied Biosystems). Each sequencing reaction was performed in a Perkin Elmer 9600 thermocycler with 25 cycles of 96°C for 10 seconds, 50°C for 5 seconds and 60°C for 2 minutes. The samples were purified using Centriflex Gel Filtration cartridges (Edge Biosystems). The samples were heated to 85°C for 2 minutes and inserted into the Prism 377 Sequencer. The sequences were analyzed using the Sequnccher™ Sequence Analysis software (Gene Codes Corp.). The sequences of the PCR product generated from human heart and human muscle were identical. The full length huE3 α I clone was obtained by ligating the two PCR products together at their XbaI site.

The nucleic acid sequence of huE3 α I (SEQ ID NO: 1), consists of an open reading frame of 5247 nucleotides which encodes a 1749 amino acid polypeptide, in addition to 695 bp in the 5' untranslated region and 362 bp in the 3' untranslated region. Alignment of the human and mouse amino acid sequence, as shown in Figure 1 (SEQ ID NOS: 2 and 15, respectively), exhibited 92.5% overall sequence identity.

In the present invention, a novel full length human E3 α cDNA (huE3 α I; SEQ ID NO: 1) was isolated and cloned and the full length polypeptide sequence (SEQ ID NO: 2) was disclosed. A partial sequence of the human E3 α gene had been previously reported. (See U.S. Patent 5,861,312; Kwon *et al.*, *Proc. Natl. Acad. of Sci. USA*, 95: 7898-7903, 1998). The reported partial sequence is encompassed in SEQ ID NO: 1; but only represents a small portion of the entire full length gene (nucleotides 702 to 1066).

EXAMPLE 2

Cloning of cDNA Encoding Human E3 α Ortholog, huE3 α II

BLAST analysis of the Amgenesis database (Amgen internal EST database) with the human E3 α I amino acid sequences revealed 4 Amgenesis EST sequences (amgi-039645, smop2-0079f12 and zhgb-aa693825 and Genebank accession no.: AA002347) which encode potential regions of the human and mouse E3 α ubiquitin ligase ortholog nucleotide sequences which are deonted as E3 α II. Based on the zhgb-aa693825 and AA002347 sequences, two PCR primer sets (#2380-88/2378-32 and #2381-48/2385-94) were designed. These sequences are set out below in Table IV.

Table IV

Primer	Sequence	SEQ ID NO:
2380-88	ATG GCG TCG CTA GAG CCA	11
2378-32	CAA AGC GGC TGA GCA TGA TCA TC	12
2381-48	TGA ACA GCC AAT CAC ACT AAG CA	13
2385-94	TTA TAA ATG CCA GTC AAT GCC AA	14

The primer sets #2380-88/2378-32 (SEQ ID NOS: 11 and 12) and #2381-48/2385-94 (SEQ ID NOS: 13 and 14) were used to generate two PCR products which encode the coding region of a novel ortholog of human E3 α ligase (huE3 α II). The 5' and 3' untranslated regions of huE3 α II were determined based on the EST sequences amgi-03645 and smop2-0079f12 in order to obtain the full length huE3 α II cDNA. PCR was performed as described above utilizing Marathon-Ready cDNA from human heart and human muscle libraries. The two PCR products were electrophoresed on a 1% gel as described by Sambrook *et al.*, *supra*. The appropriate sized bands (2.2 kB and 3.5 kB) were excised from the agarose and purified by QIAquick Gel Extraction kit (cat.# 28704). The PCR products were subcloned into the pcDNA3.1-HisA vector (Invitrogen cat.# V385-20) and transformed into *E.coli* (Strain INV α F) using the Invitrogen Original TA Cloning kit. The insert DNA was purified with the QIAprep Spin Miniprep kit (QIAGEN cat.# 27104) and subsequently digested with NotI/SacI for the 2 kB product and SacI/XhoI for the 3.3 kB product. The PCR products were sequenced as described in Example 1 and the products generated from human heart and human muscle cDNA libraries were identical. The full length huE3 α II gene was generated by ligating these two PCR products at their SacI sites.

The nucleic acid sequence of huE3 α II (SEQ ID NO: 3), consists of an open reading frame of 5265 nucleotides which encodes a 1755 amino acid polypeptide, in addition to 294 bp in the 5' untranslated region and 740 bp in the 3' untranslated region. Alignment of the human and mouse amino acid E3 α II sequences, as shown in Figure 1 (SEQ ID NOS: 4 and 6, respectively), exhibited 90.4% overall sequence identity. There is a 48.1% overall amino acid sequence identity between human E3 α I and human E3 α II (SEQ ID NOS:2 and 4, respectively).

In the present invention, a novel full length cDNA sequence encoding huE3 α II (SEQ ID NO: 3) was cloned and isolated, and the full length polypeptide sequence was disclosed (SEQ ID NO: 4). A partial sequence of huE3 α II was identified in WO9904265 as one of many partial sequences with unknown identities that were speculated to be cancer markers.

EXAMPLE 3

Cloning of the Murine E3 α II Ortholog

BLAST analysis of the Amgen internal database, Amgenis, with human E3 α II amino acid sequences identified the mouse cDNA clone (Smop2-00079-f12) as a potential mouse ortholog of E3 α II ubiquitin ligase. The Amgenis database contained the entire coding region of the mouse E3 α II ubiquitin ligase (muE3 α II) gene. The cDNA clone of was obtained from the Amgen sequencing group. The sequence of the clone was confirmed to be the full cDNA of muE3 α II as described in Example 1. The nucleic acid sequence of muE3 α II (SEQ ID NO: 5), consists of an open reading frame of 5265 nucleotides which encodes a 1755 amino acid polypeptide, in addition to 765 bp in the 5' untranslated region and 56 bp in the 3' untranslated region.

EXAMPLE 4

Human E3 α II Tissue Expression

Tissue expression patterns of huE3 α I and huE3 α II mRNA were analyzed by Northern blot analysis. To detect the presence of huE3 α II transcript in various tissues, a ³²P-labeled fragment of huE3 α II, which was 452 bp and corresponded to nucleotides 3557-4009 of SEQ ID NO: 3, was used as a probe. For detection of huE3 α I transcript in various tissues, a ³²P-labeled fragment of huE3 α I, which was 696 bp and corresponded to nucleotides 3468 -4164 of SEQ ID NO: 1, was used as a probe.

The probes were labeled by random priming method using Prime-it RMT labeling kit (Stratagene, Cat# 300392). The specific activities was 1.436×10^6 cpm/ μ l for the huE3 α II probe and 1.207×10^6 cpm/ μ l for the huE3 α I probe. Human multiple tissue poly A+ RNA blots (Clontech cat.# 7780-1) were prehybridized in Church hybridization solution (1% BSA, 7% SDS, 0.5 M sodium phosphate, pH 7.0, 1 mM EDTA) for 4 hour at 65°C. The blots are then hybridized in Church hybridization solution with 3.0×10^6 cpm/ml 32 P labeled probe for overnight at 65°C. The blots are then washed 3 times in Wash B buffer (1% SDS, 0.04 M sodium phosphate, 1 mM EDTA) for 5 minutes each at room temperature, followed by two times at 65°C. The blots were exposed to X-ray film at room temperature overnight (for huE3 α II detection) or one week (for huE3 α I detection).

The Northern blot analysis revealed that huE3 α II (Figure 2) is predominantly expressed in skeletal muscle, with moderate expression in heart and kidney tissue and minimal or no expression in other tissues examined including brain, colon, thymus, spleen, liver, small intestines, placenta, lung and peripheral white blood cells. In contrast, the expression of huE3 α I (Figure 3) is less muscle-specific. Although heart and skeletal muscle had relative high levels of huE3 α I transcripts, moderate levels of huE3 α I was found to spread through the various tissues examined. The results indicate that huE3 α II is the more muscle-specific form of huE3 α which is predominantly expressed in skeletal muscle tissue.

EXAMPLE 5

Production of huE3 α Polypeptides

A. Bacterial Expression of huE3 α Polypeptides

PCR is used to amplify template DNA sequences encoding a huE3 α polypeptide using primers corresponding to the 5' and 3' ends of the sequence. The amplified DNA products may be modified to contain restriction enzyme sites to allow for insertion into expression vectors. PCR products are gel purified and inserted into expression vectors using standard recombinant DNA methodology. An exemplary vector, such as pAMG21 (ATCC No. 98113) containing the *lux* promoter and a gene encoding kanamycin resistance is digested with BamHI and NdeI for directional cloning of inserted DNA. The ligated mixture is transformed into an *E. coli* host strain by electroporation and transformants are selected for kanamycin resistance. Plasmid DNA

from selected colonies is isolated and subjected to DNA sequencing to confirm the presence of the insert.

Transformed host cells are incubated in 2x YT medium containing 30 µg/ml kanamycin at 30°C prior to induction. Gene expression is induced by the addition of N-(3-oxohexanoyl)-dl-homoserine lactone to a final concentration of 30 ng/ml followed by incubation at either 30°C or 37°C for six hours. The expression of huE3α polypeptide is evaluated by centrifugation of the culture, resuspension and lysis of the bacterial pellets, and analysis of host cell proteins by SDS-polyacrylamide gel electrophoresis.

Inclusion bodies containing huE3α polypeptide are purified as follows. Bacterial cells are pelleted by centrifugation and resuspended in water. The cell suspension is lysed by sonication and pelleted by centrifugation at 195,000 x g for 5 to 10 minutes. The supernatant is discarded, and the pellet is washed and transferred to a homogenizer. The pellet is homogenized in 5 ml of a Percoll solution (75% liquid Percoll/0.15 M NaCl) until uniformly suspended and then diluted and centrifuged at 21,600 x g for 30 minutes. Gradient fractions containing the inclusion bodies are recovered and pooled. The isolated inclusion bodies are analyzed by SDS-PAGE. A single band on an SDS polyacrylamide gel corresponding to *E. coli*-produced huE3α polypeptide is excised from the gel, and the N-terminal amino acid sequence is determined essentially as described by Matsudaira *et al.*, *J. Biol. Chem.*, 262: 10-35 (1987).

B. Mammalian Cell Production of huE3α Polypeptides

The huE3α DNA was subcloned into a mammalian expression vector as described above using standard DNA technology. An exemplary expression vector, pCEP4 (Invitrogen, Carlsbad, CA), which contains an Epstein-Barr virus origin of replication, may be used for the expression of huE3α in 293-EBNA-1 cells. Amplified and gel purified PCR products are ligated into pCEP4 vector and lipofected into 293-EBNA cells. The transfected cells are selected in 100 µg/ml hygromycin and the resulting drug-resistant cultures are grown to confluence. The cells are then cultured in serum-free media for 72 hours. The conditioned media is removed and, huE3α protein polypeptide expression is analyzed by SDS-PAGE. Human E3α polypeptide expression may be detected by silver staining. Alternatively, huE3α polypeptide is produced as a

fusion protein with an epitope tag, such as an IgG constant domain or a FLAG epitope, which may be detected by Western blot analysis using antibodies to the tag peptide.

Human E3 α polypeptides may be excised from an SDS-polyacrylamide gel, or huE3 α fusion proteins are purified by affinity chromatography to the epitope tag, and subjected to N-terminal amino acid sequence analysis as described herein.

EXAMPLE 6

Production of Anti-huE3 α Polypeptide Antibodies

Antibodies to huE3 α polypeptides may be obtained by immunization with purified protein or with huE3 α peptides produced by biological or chemical synthesis. Suitable procedures for generating antibodies include those described in Hudson and Bay, *Practical Immunology, Second Edition*, Edition, Blackwell Scientific Publications.

In one procedure for the production of antibodies, animals (typically mice or rabbits) are injected with a huE3 α antigen (such as a recombinant truncated form of huE3 α polypeptide), and those with sufficient serum titer levels as determined by ELISA are selected for hybridoma production. Spleens of immunized animals are collected and prepared as single cell suspensions from which splenocytes are recovered. The splenocytes are fused to mouse myeloma cells (such as Sp2/0-Ag14 cells), allowed to incubate in DMEM with 200 U/ml penicillin, 200 μ g/ml streptomycin sulfate, and 4 mM glutamine, then incubated in HAT selection medium (Hypoxanthine; Aminopterin; Thymidine). After selection, the tissue culture supernatants are taken from each fusion well and tested for anti-huE3 α antibody production by ELISA.

Alternative procedures for obtaining anti-huE3 α antibodies may also be employed, such as the immunization of transgenic mice harboring human Ig loci for production of human antibodies, and the screening of synthetic antibody libraries, such as those generated by mutagenesis of an antibody variable domain.

EXAMPLE 7

Biological Activity of huE3 α polypeptides

Human E3 α family members are known to catalyze the ubiquitin conjugation reaction which ultimately results in protein degradation. To determine the biological activity of huE3 α polypeptide, the rate of ubiquitin conjugation and the

rate of protein degradation are measured. The following are examples of assays to measure these biological activities.

A. Ubiquitin Conjugation Assay:

The enzymatic activity of E3 α family members is thought to be the rate limiting step in ubiquitin conjugation. Rat skeletal muscles are dissected, homogenized, and centrifuged at 100,000 x g to remove proteosomes. The soluble extract is incubated with ¹²⁵I-ubiquitin (Amersham, Arlington Heights, IL) (0.15 mg/ml) in 20 mM Tris (pH 7.4), 1 mM DTT, 5 mM MgCl₂, and 2 mM ATP γ S at 37°C in the presence and absence of huE3 α polypeptide. At various time points, the reactions are terminated by the addition of sample buffer and SDS-PAGE is performed on a 12% gel. The gel is then dried and autoradiographed. If huE3 α acts as an E3 α family member, the level of ubiquitination should increase in extracts treated with the huE3 α polypeptide (Soloman *et al.*, *Proc. Natl. Acad. of Sci. U.S.A.*, 95: 12602-07, 1998).

B. Protein Degradation Assays:

Measurement of tyrosine release is a preferred method for determining the rate of protein turnover in skeletal muscles. Rat skeletal muscles are dissected and homogenated. The extracts are incubated at 37°C for 2 hours in 20 mM Tris (pH 7.6), 5 mM MgCl₂, 2 mM DTT, ATP-regenerating system (10 μ g creatine phosphokinase and 10 mM creatine phosphate), 1 mM ATP, and 25 mg of ubiquitin in the presence and absence of huE3 α polypeptide. Subsequently, the reactions are terminated with 20% TCA. After centrifugation, the concentrations of tyrosine in the supernatant is measured by fluorescence spectroscopy according to the method of Waakkes and Udenfriend (*J. Lab. Clin. Med.*, 50: 733-736, 1957).

Measurement of radiolabeled proteins will also indicate if huE3 α polypeptide exhibits E3 α family biological activity. Rat skeletal muscle homogenates are incubated at 37°C for 2 hours with ¹²⁵I-labeled N-end pathway substrates, such as ¹²⁵I-lysozyme and ¹²⁵I-lactalbumin, in the presence and absence of huE3 α polypeptide.

Following the incubation, 20% TCA is added to precipitate the radioactivity. The release of TCA-soluble radioactivity is measured using a gamma counter and correlates the rate of protein degradation. The addition of huE3 α polypeptide should increase the rate of protein degradation in both of these assays.

EXAMPLE 8

Identification of Modulators of the Biological Activity of huE3 α Polypeptides

The assays described in Example 7 demonstrate preferred methods to measure the biological activity of huE3 α as an ubiquitin ligase. These methods are also useful for identifying modulators of huE3 α ubiquitin ligase activity.

The rate limiting step of ubiquitin conjugation consists of E3 α catalyzing the transfer of the activated ubiquitin molecule to the target protein. The rate of ubiquitination modulated by huE3 α can be measured in dissected rat skeletal muscles as described in Example 7. The addition of potential huE3 α modulators (inhibitors or stimulators) to this system will allow for the identification of E3 α stimulators and inhibitors by virtue of their ability to modulate the level or rate of ubiquitin conjugation to the target protein. If the addition of the modulator decreases the rate of huE3 α -modulated ubiquitin conjugation, it is considered a huE3 α inhibitor. If the modulator increases the rate of ubiquitin conjugation it is considered a stimulator.

The effect of huE3 α modulators can also be determined by measuring their effect on the rate of protein turnover as described in Example 7. If huE3 α exhibits the biological activity of an ubiquitin ligase, it will induce protein degradation. Protein turnover is measured by quantitating tyrosine release or the degradation of radioactively labeled N-end pathway substrates in the presence of E3 α modulators. The addition of effective huE3 α modulators will either increase or decrease the rate of protein degradation.

EXAMPLE 9

Identification of huE3 α I Single Nucleotide Polymorphisms (SNP)

A BLAST search of the Celera Human Genome database was conducted using the huE3 α I cDNA sequence (SEQ ID NO: 1) as a probe. The sequences identified in the search were used to manually assemble a polynucleotide sequence (SEQ ID NO: 18) which was discovered to have a single nucleotide mismatch at nucleotide 5397 of the huE3 α I cDNA sequence (SEQ ID NO: 1). The polynucleotide sequence of SEQ ID NO: 18 contains a huE3 α I SNP with a change of a cytosine to a thymidine at position 4702, which caused the predicted amino acid sequence of SEQ ID NO: 2 to change from an Arg residue to a W (Trp) residue at position 1508.

PCR was carried out to confirm the polynucleotide sequence of huE3 α I cDNA. Primers were designed to flank the mismatch as follows: 5' AGAAGGAGAGTACAGTGCACTC3' (SEQ ID NO: 20) and 5'CGAAAGCATCCTGTCCTCTG (SEQ ID NO: 21). PCR was carried out as described in Example 1 with the Marathon-Ready cDNA library (Clontech cat no. 7413-1) from which huE3 α I cDNA was cloned. The PCR reactions resulted in 8 individual PCR products which had identical sequences to the huE3 α I SNP (SEQ ID NO: 18).

These experiments have confirmed the sequence of a huE3 α I SNP set out in SEQ ID NO: 17 wherein the nucleotide at position 4702 is a thymidine. Accordingly, the correct predicted amino acid sequence is set out as SEQ ID NO: 19, wherein the residue at position 1568 is W (Trp).

EXAMPLE 10

Human E3 α I and E3 α II Stimulate Ubiquitination

To confirm that huE3 α I and huE3 α II have the predicted enzymatic activity of stimulating ubiquitin conjugation, ubiquitination reactions were carried out in 293 cells. Cultures of 293T cells (ATCC accession no. CRL1573) were transfected with huE3 α I or huE3 α II full length cDNA (SEQ ID NOS: 1 or 3, respectively) that had been subcloned into pcDNA3.1 vector (Invitrogen) under the control of the CMV promoter using Lipofectamine reagent 2000 (Gibco, cat no. 11668-027) according to the manufacture's instructions. As a control, 293T cultures were transfected with pcDNA3.1 vector without the cDNA insert. The transfected cells were lysed in ice-cold lysis buffer (50mM Tris-HCl (pH 8.0), 2mM DTT, 5 mM MgCl₂) in the presence of Sigma P8340 protease inhibitor cocktail (containing 4-(2-aminoethyl) benzenesulfonyl fluoride, pepstatin A, E-64, bestatin, leupeptin and aprotinin) at 100 μ l/10⁷ cells. The crude lysates were then centrifuged at 10,000 g for 10 minutes.

The supernatants prepared from vector- (Control), human E3 α -I- (hu-E3 α -I) or human E3 α -II- (hu-E3 α -II) transfected cells were subjected to ubiquitination reactions. To measure ubiquitination of endogenous proteins, 30 μ g of cell lysate was incubated with ¹²⁵I-ubiquitin (0.15 mg/ml, approximately 10⁷cpm) in a total volume of 40 μ l in a buffer containing 50 mM Tris, pH 8.0, 2 mM DTT, 5 mM MgCl₂, 2 mM adenosine 5'-[thio]triphosphate (ATP S), 50 μ g/ml ubiquitin

aldehyde, MG132 20 µg/ml and protease inhibitor cocktail (Sigma P8340) at 37°C for 30 minutes. Reactions were stopped by adding sample buffer and were subjected to 12% SDS PAGE. The gels were then dried and autoradiographed.

The ubiquitination of α -lactalbumin, a known substrate for N-end Rule Ubiquitination was also measured with the 239T transfected cells. For these reactions, 30 µg of cell lysate proteins was incubated with 0.15 mg/ml ¹²⁵I- α -Lactalbumin and 0.25 mg/ml unlabeled ubiquitin in a total volume of 40 µl in a buffer containing 50 mM Tris (pH 8.0) 2 mM DTT, 5 mM MgCl₂, 2 mM adenosine 5'-[thio]triphosphate (ATP S), ubiquitin aldehyde 50 ug/ml, 20 µg/ml MG132 and protease inhibitor cocktail (Sigma P8340) at 37°C for 30 minutes. Reactions were stopped by adding sample buffer and each reaction was run of a 8% SDS PAGE was performed. The gels were then dried and autoradiographed

The amount of radioactivity incorporated into high molecular weight bands denoted as "ubiquitin-protein conjugates" in Figure 4 (above 18 kDa for endogenous proteins and above 35 kDa for α -Lactalbumin) were quantitated by using PhosphorImager and plotted (right panel). These reactions indicated that recombinant expression of huE3 α I or huE3 α II in 293 cells lead to accelerated ubiquitination of endogenous cellular proteins and ubiquitin conjugation to α -lactobumin, a bona fide N-end rule substrate.

To further substantiate the enzymatic activity of huE3 α I and huE3 α II, ubiquitin conjugation to endogenous cellular proteins were measured in cultured muscle cell lines. Cultures of murine C₂C₁₂ or rat L6 myotube cells (ATCC accession nos. CRL-1772 and CRL-1458, respectively) were transfected with huE3 α I or huE3 α II full length cDNA under control of the CMV promoter using Lipofectamine 2000 Reagent (Gibco). Mock transfection with the pcDNA3.1 vector without a cDNA insert was performed as a control. Cell lysates were prepared as described above for the 293T cells and the resulting supernatants were used in ubiquitin conjugation reactions. For each reaction, 30 µg of C₂C₁₂ or L6 myotube cell lysate was incubated with ¹²⁵I-ubiquitin (0.15 mg/ml, approximately x10⁷cpm) in a total volume of 25-30 µl in a buffer containing 50 mM Tris, pH 8.0, 2 mM DTT, 5 mM MgCl₂, 2 mM adenosine 5'-[thio]triphosphate (ATP S), 50 µg/ml ubiquitin aldehyde, 20 µg/ml MG132 and protease inhibitor cocktail (Sigma P8340) at 37°C for 30 minutes. Reactions were stopped by adding sample buffer and were subjected to 12% SDS PAGE. The gels were then dried and autoradiographed.

The amount of ubiquitinated muscle proteins (^{125}I -Ubiquitin protein conjugates) were quantitated as the total radioactivity incorporated into high molecular weight bands (above 18 kDa) using a Phosphorimager as shown in Figure 5 (left panel). These reactions indicated that transfection of huE3 α I and huE3 α II increased ubiquitination of cellular proteins 2-3 fold (see Figure 5, right panel) in murine C₂C₁₂ and rat myotube cultures.

EXAMPLE 11

Expression of Human E3 α I and Human E3 α II is Unregulated

During Cachexia Disease States

The Yoshida Hepatoma-130 (YAH) cachexia rat model as described in Baracos *et al.* (*Am. J. Physiol.*, 268(5 Pt 1): E996-1006, 1995) was used to determine if huE3 α I and huE3 α II are upregulated in cachexia disease states. For tumor implantation, female Sprague-Dawley rats of the Buffalo strain from a colony maintained at the University of Alberta were used as the host for the YAH tumor cells. Tumor cell stocks were maintained in liquid nitrogen and used after two passages in recipient female animals of the same strain. Rats were housed in individual wire mesh cages in a temperature (24°C)- and humidity (80%)-controlled room on a 12:12-h light-dark cycle. Rats were fed ground laboratory chow (Continental Grain, Chicago, IL) containing 24% crude protein.

Rats were allocated by initial body weight to three groups such that the sizes (mean \pm SE) of the animals receiving each treatment were similar (~ 200g). Two different treatments were compared: YAH-bearing and pair-fed control rats. The pair-fed rats, which received one meal per every day at 9.00am, were fed on the basis of their body weights, the same amount of food consumed by the tumor-bearing rats. On days 1,2,3,4 and 5 after tumor-implantation, food intake was determined in preliminary experiments to be 9, 7.5, 5.3, 1.5, and 0.9% respectively, of initial body weight per day. Rats were implanted with 100 ml of ascites fluid containing YAH cells from a single donor animal. The control rats were implanted with an equal volume of saline buffer. Rats were sacrificed by CO₂ asphyxiation after 3 and 5 days, and epitrochlearis, EDL, soleus, medial gastrocnemius muscles were rapidly dissected and the gastrocnemius muscles were weighed. Tissues were frozen immediately in liquid nitrogen and stored at -70°C until use.

The gastrocnemius skeletal muscle weights in YAH-130 tumor bearing rats were significantly lower than those measured from the pair-fed control rats. As indicated in the Table V below, the YAH-130 tumor bearing rats underwent muscle wasting by day 3 after tumor implantation which was more apparent at day 5 after implantation. The muscle weights are calculated (in grams) as the mean \pm standard error.

Days after Tumor Implantation	n	Pair Fed Control (in grams)	Tumor-Bearing (in grams)	Percent Change
3 days	8	530 \pm 14.6	508 \pm 7.3	-4.3%
5 days	8	593 \pm 8.1	443 \pm 9.4	-25.3%

n = number of animals

The rate of ubiquitin conjugation of the endogenous muscle proteins were carried out as described in Example 10 using the skeletal muscles from the YAH tumor-bearing rats. The frozen gastrocnemius muscles collected (via dissection at sacrifice) from 6 tumor-bearing rats were combined. The muscle extracts (20% weight/volume) were prepared by homogenizing the muscles in a buffer containing 50 mM Tris HCl (pH 8.00), 5 mM MgCl₂, 2 mM DTT, protease inhibitor cocktail (Sigma P8340) and 10% glycerol. The homogenates were then centrifuged at 40,000g for 1 hr and the resulting supernatants were used as crude muscle extracts.

For some assays, the crude muscle extracts were fractionated further by chromatography on DEAE-cellulose (Whatman, Clifton, New Jersey) to remove endogenous ubiquitin as described by Solomon *et al.* (*Proc. natl. Acad. Sci. U.S.A.*, 95: 12602-7, 1998). The bound material Fraction II, which contained most of the ubiquitin conjugating enzymes were eluted with 50 mM Tris, pH 8.0 containing 0.5M NaCl and 1 mM DTT. Both crude extracts and Fraction II were dialyzed prior to use for ubiquitination assay against buffer containing 20 mM Tris, pH 8.0, 2 mM DTT, 5 mM MgCl₂, and 10% glycerol and stored at 70 °C until use. Crude muscle extracts were used for ubiquitin conjugation to 125I- α -lactalbumin. Fraction II was used when rates of endogenous skeletal muscles proteins were compared and also when effects of E3 α inhibitors on skeletal muscle protein ubiquitination were tested.

The Fraction II from both tumor-bearing and pair-fed control rats were subjected to ubiquitination reactions of the endogenous muscle proteins as described in

Example 10 in the presence of 20 µg/ml of bestatin and 10 mM of either the E3α selective inhibitor arginine methyl ester (Arg-ME) or the control alanine methyl ester (Ala-ME) (Sigma Chemicals, St. Louis MO). The reactions were incubated at 37°C for 20 minutes and the ¹²⁵I-Ubiquitin conjugates were resolved by 12% SDS PAGE as described in Example 10.

As shown in Figure 6, the tumor-bearing rats exhibited accelerated muscle protein ubiquitination. The increase in ubiquitination within the rat skeletal muscles of the tumor-bearing rats was attributable to the activation of the E3α/N-end rule pathway, since the addition of E3α specific inhibitor arginine methylester virtually abolished the accelerated ubiquitination activity (see lanes 9 and 10 on Figure 6).

To further establish the role of huE3αI and huE3αII in the N-end rule pathway in muscle wasting in the rat cachexia model, the rates of ubiquitination of N-end rule substrate α-lactalbumin was measured in skeletal muscle extracts from control and tumor-bearing mice. ¹²⁵I-α-lactalbumin (0.15 mg/ml) was incubated with crude skeletal muscle extracts (2 mg/ml) in the presence of 0.25 mg/ml of ubiquitin at 37°C for 0 or 20 minutes as described in Example 10. As shown in Figure 7, the atrophying muscles dissected from the tumor-bearing rats exhibited increased ubiquitin conjugation to ¹²⁵I-α-lactalbumin

Northern blot analysis was carried out to measure the huE3αI and huE3αII mRNA expression in the gastrocnemius muscles of YAH-130 tumor-bearing mice. RNA from the dissected muscles was isolated with Trizol Reagent (Gibco, cat: 15596-018). The final RNA pellets were resuspended in DEPC-H₂O and 20 µg of total RNA per lane were separated by electrophoresis through 1% agarose gels. The separated RNA was transferred to nylon membranes and cross-linked to the filter by exposure to ultraviolet light.

The cDNA probes were generated by PCR with the following primers: for Hu-E3α-I probe: 5' primer, AGG AAG CTG TGG TCA TGT (SEQ ID NO: 22) ; 3' primer, GTT AGG AAG AAC AAC TG (SEQ ID NO: 23); for Hu-E3α-II probe: CTA AAG AAC AGC GAA GGC AAC AG (SEQ ID NO: 24); 3' primer, CGC AGC TAC CCC AAC ACA TTA T (SEQ ID NO: 25). PCR was carried out for 30 cycles at 94°C for 45 seconds, 50-58 °C for 45 seconds, and 72°C for 1 minutes using a commercially available kit (Boehringer Mannheim, cat: 1578553,). The PCR product was cloned into the pCR2.1 vector using the Original TA Cloning kit (Invitrogen). After digestion with EcoRI, the cloned PCR product was sequenced and confirmed.

The resulting cDNA probes were radiolabeled with [³²P]dCTP using the Prime-it RMT labeling kit (Stratagene, cat: 300392). Membranes were prehybridized and hybridized (with the cDNA probes) in buffer containing 1% BSA, 7% SDS, 0.5 M Sodium Phosphate (pH 7.0), 1mM EDTA. Subsequently, the blots were washed in
5 buffer containing 1% SDS, 0.04 M Sodium Phosphate, 1 M EDTA, by the method of Church and Gilbert (*Proc. Natl. Acad. Sci. U.S.A.* 81:1991-1995, 1984) and exposed to radiographic film at -70°C overnight.

As shown in Figure 8, the expression of huE3αII mRNA was increased at day 3 post-tumor implantation (see left panel) but the level of huE3αI had not changed
10 significantly at day 3 (see right panel). This coincides with the time point when the significant decrease in muscle mass was detectable in the C26 tumor-bearing mice (See Table VI in Example 12). The expression of both huE3αI and huE3αII was elevated at day 5 post-tumor implantation in the tumor bearing rats. This corresponds to a cachexia state with severe muscle wasting.

EXAMPLE 12

Expression of Human E3αI and Human E3αII in a Murine Cancer Cachexia Model of C26 Tumor Bearing Mice

The Colon-26 (C-26) tumor model of cachexia was used to demonstrate the role of huE3αI and huE3αII as described in Matsumoto *et al.*, *Brit. J. Can.* 79: 764-9 (1999) and Tanaka *et al.*, *Can. Res.*, 50: 2290-5 (1990). Seventy-two week old male CDF1 mice were injected in the left flank with 0.2 ml containing
25 either 0.5 x10⁶ C26 cells or PBS. Following injection, body weight and food intake was observed daily. The pair fed control mice (generated as described in Example 11) were fed the daily average food intake of the tumor bearing group. On the day 12 or 17 post-injection, tumor bearing mice and pair fed control mice were sacrificed by CO₂ asphyxiation. Subsequently, a terminal serum sample was collected and the
30 kidney, heart and gastrocnemius muscles were rapidly dissected and weighed. The resulting C26 tumors were also weighed. The tissues were frozen on dry ice and stored at -70°C.

The wet weight of the skeletal muscles from the tumor-bearing mice were significantly less than the weight of those from the pair-fed control mice as
35 shown below in Table VI:

Days after Tumor Implantation	n	Pair Fed Control (in grams)	Tumor-Bearing (in grams)	Percent Change
12 days	12	0.127 ± 0.007	0.116 ± 0.0072	-8.6%
17 days	12	0.117 ± 0.009	0.087 ± 0.001	-26%

n = number of animals

RNA was isolated at day 12 and day 17 from the gastrocnemius muscle and cardiac muscle from the C26 tumor-bearing and pair-fed control mice as described in Example 11. Northern blot analysis was carried out by loading 20 µg of total RNA per lane and separating by electrophoresis through 1% agarose gels. The separated RNA was transferred to nylon membranes and cross-linked to the filter by exposure to ultraviolet light.

The cDNA probes were generated by PCR with the following primers: for mouse E3α-I probe: 5' primer, TTT CTT CCA TTC CCT GCA TAC A (SEQ ID NO: 26), 3' primer, CAA AAC TTT ATA AAG GTG CCC GTA A (SEQ ID NO: 27), and for Mouse E3α-II probe: 5' primer, ATT CCC TGC ATG CAC TTC AGT AA (SEQ ID NO: 28), 3' primer, CAT TCC CTG CAT GCA CTT CAG SEQ ID NO: 29). PCR was carried out for 30 cycles at 94°C for 45 seconds, 50-58 °C for 45 seconds, and 72°C for 1 minutes using a commercially available kit (Boehringer Mannheim cat: 1578553,). The PCR product was cloned into the pCR2.1 vector using the Original TA Cloning kit (Invitrogen). After digestion with EcoRI, the cloned PCR product was sequenced and confirmed.

The resulting cDNA probes were radiolabeled with [³²P]dCTP using the Prime-it RMT labeling kit (Stratagene, cat: 300392). Membranes were prehybridized and hybridized (with the cDNA probes) in buffer containing 1% BSA, 7% SDS, 0.5 M Sodium Phosphate (pH 7.0), 1mM EDTA. Subsequently, the blots were washed in buffer containing 1% SDS, 0.04 M Sodium Phosphate, 1 M EDTA, by the method of Church and Gilbert (*Proc. Natl. Acad. Sci. U.S.A.* 81:1991-1995, 1984) and exposed to radiographic film at -70°C overnight.

As shown in Figure 9, at day 12 after tumor-implantation there was a clear increase in huE3αII mRNA expression in the skeletal muscles of tumor-bearing mice. Expression of both huE3αI and huE3αII was increased at day 17 post-implantation. Increased expression of huE3αII mRNA coincides with the time point

when the significant decrease in muscle mass became detectable in tumor-bearing mice (See Table VI above). The expression of huE3 α I and huE3 α II remained unchanged in the cardiac muscle of the tumor-bearing mice. This corresponds to a cachexia state with severe muscle wasting.

The data described in both Examples 11 and 12 show that in experimental cachexia models, there was a sharp rise in the rate of ubiquitination in skeletal muscle tissues. The accelerated ubiquitination is due largely to the activation of E3 α , since addition of the E3 α -selective inhibitor, arginine methylester, virtually abolished all the increased ubiquitination activities. In addition, the data demonstrated that in two widely used experimental models of cachexia (murine C26 tumor-bearing model and rat YAH-130 tumor-bearing model), the mRNA levels of E3 α -I and E3 α -II increase significantly and specifically within skeletal muscle during the course of cachexia and muscle wasting. In these disease models and during the course of cachexia, the induction of E3 α -II occurred earlier than that of E3 α -I and coincided with the early onset of muscle wasting. During the late stage of cachexia, both E3 α -I and E3 α -II were markedly induced when muscle wasting became pronounced. Therefore, the results suggest that E3 α -II may play a more critical role in cachexia, although both E3 α -I and E3 α -II are apparently involved in the disease process.

EXAMPLE 13

Treatment of Muscle Cells with TNF α and IL-6 leads to Increased Expression of Human E3 α II and Increased Ubiquitination

Treatment with the proinflammatory cytokines, TNF α and IL-6, caused the induction of huE3 α II in C₂C₁₂ myotube cultures. C₂C₁₂ myoblasts were cultured in 100-mm dishes in an atmosphere of 5% CO₂ at 37°C in DMEM supplemented with 10% FBS and L-glutamine to reach 100% confluence. Myoblast differentiation was induced with DMEM supplemented with 2% horse serum and L-glutamine for 96 hours. Differentiated myotubes were then treated with TNF α (10 ng/ml; R&D Systems cat no. 210-TA) or IL-6 (10 ng/ml; R&D Systems cat no. 206-IL) for 3 days and 5 days.

After the 3 or 5 day incubation, RNA from differentiated C₂C₁₂ cultures was isolated with Trizol Reagent and Northern blot analysis was carried out

as described in Example 11. Isolated RNA from untreated C₂C₁₂ cultures were used as a control. The blots were hybridized with a ³²P-labeled cDNA probes specific for muE3αI (lower panels) and muE3αII (upper panels). The probes were genreated as described in Example 12.

As shown in Figure 10, the expression of muE3αII was markedly increased in the cells treated with TNFα or IL-6 (See upper panels). Conversely, the expression of muE3αI was not drastically induced in response to proinflammatory cytokine treatment. This data indicates a role for E3αII in cytokine-mediated protein catabolism and muscle wasting.

Cytokine treatment also resulted in accelerated ubiquitination in differentiated C₂C₁₂ cells. C₂C₁₂ cells were differentiated for 5 days to allow formation of myotubes. The differentiated myotubes were treated with 2 ng/ml of IL-6 for 5, 24 or 48 hours. After the incubation, the cells were lysed and ¹²⁵I-Ubiquitin conjugation was carried out as described in Example 10. As shown in Figure 11, IL-6 treatment resulted in a marked increase in ubiquitination of cellular proteins (left panel) which was detectable 5 hours post-treatment. The increase in ubiquitination was time dependent (see right panel).

Differentiated C₂C₁₂ myotubes were also treated with increasing concentrations of TNFα (0, 3, 6, 10, 20 ng/ml) for one hour. This treatment resulted in a dose dependent increase in ¹²⁵I- ubiquitin conjugation of cellular proteins as shown in Figure 12.

TNFα and IL-6 are major proinflammatory cytokines known to be involved in cachexia and tissue wasting. The data reveals that these cytokines significantly upregulate the mRNA expression of E3αII in muscle cells and stimulate muscle protein ubiquitination. Proinflammatory cytokines, such as TNFα, IL-6, IL-1, interferon-gamma, CNTF and leptin, have been shown to be involved in disease states of cachexia and protein/tissue wasting, including cancer cachexia, renal cachexia (energy-protein malnutrition), burn cachexia and AIDS wasting. These findings that TNFα and IL-6 induce the expression of E3αII (Figure 9) and stimulate protein ubiquitination in muscle cells (Figures 10 and 11) strongly suggest that E3αII is critical target via which various cachectic factors induce protein catabolism and cachexia/muscle wasting. This argument is further supported by our finding that

recombinant expression of E3 α II by cDNA transfection leads to marked protein ubiquitination in myotube cultures (Figure 5).

While the present invention has been described in terms of the preferred embodiments, it is understood that variations and modifications will occur to those skilled in the art. Therefore, it is intended that the appended claims cover all such equivalent variations which come within the scope of the invention as claimed.

CLAIMS

WHAT IS CLAIMED IS:

1. An isolated nucleic acid molecule comprising a nucleotide
5 sequence selected from the group consisting of:
 - (a) the nucleotide sequence as set forth in SEQ ID NOS: 1 or 3;
 - (b) a nucleotide sequence encoding the polypeptide set forth in
SEQ ID NOS: 2 or 4;
 - (c) a nucleotide sequence which hybridizes under moderately or
10 highly stringent conditions to the complement of (a) or (b); and
 - (e) a nucleotide sequence complementary to any of (a)-(c).
2. An isolated nucleic acid molecule comprising a nucleotide
sequence selected from the group consisting of:
 - 15 (a) a nucleotide sequence encoding a polypeptide that is at least
about 70, 75, 80, 85, 90, 95, 96, 97, 98, or 99 percent identical to the polypeptide set
forth in SEQ ID NOS: 2 or 4, wherein the encoded polypeptide has an activity of the
polypeptide set forth in SEQ ID NOS: 2 or 4;
 - (b) a nucleotide sequence encoding an allelic variant or splice
20 variant of the nucleotide sequence as set forth in SEQ ID NOS: 1 or 3;
 - (c) a nucleotide sequence of SEQ ID NOS: 1; 3; (a); or (b)
encoding a polypeptide fragment of at least about 25 amino acid residues, wherein the
polypeptide has an activity of the polypeptide set forth in SEQ ID NOS: 2 or 4;
 - (d) a nucleotide sequence of SEQ ID NOS: 1, 3, or (a)-(c)
25 comprising a fragment of at least about 16 nucleotides;
 - (e) a nucleotide sequence which hybridizes under moderately or
highly stringent conditions to the complement of any of (a)-(d); and
 - (f) a nucleotide sequence complementary to any of (a)-(c).

3. An isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence encoding a polypeptide set forth in SEQ ID NOS: 2 or 4 with at least one conservative amino acid substitution, wherein the polypeptide has an activity of the polypeptide set forth in SEQ ID NOS: 2 or 4;

(b) a nucleotide sequence encoding a polypeptide set forth in SEQ ID NOS: 2 or 4 with at least one amino acid insertion, wherein the polypeptide has an activity of the polypeptide set forth in SEQ ID NOS: 2 or 4;

(c) a nucleotide sequence encoding a polypeptide set forth in SEQ ID NOS: 2 or 4 with at least one amino acid deletion, wherein the polypeptide has an activity of the polypeptide set forth in SEQ ID NOS: 2 or 4;

(d) a nucleotide sequence encoding a polypeptide set forth in SEQ ID NOS: 2 or 4 which has a C- and/or N- terminal truncation, wherein the polypeptide has an activity of the polypeptide set forth in SEQ ID NOS: 2 or 4;

(e) a nucleotide sequence encoding a polypeptide set forth in SEQ ID NOS: 2 or 4 with at least one modification selected from the group consisting of amino acid substitutions, amino acid insertions, amino acid deletions, C-terminal truncation, and N-terminal truncation, wherein the polypeptide has an activity of the polypeptide set forth in SEQ ID NOS: 2 or 4;

(f) a nucleotide sequence of (a)-(e) comprising a fragment of at least about 16 nucleotides;

(g) a nucleotide sequence which hybridizes under moderately or highly stringent conditions to the complement of any of (a)-(f); and

(h) a nucleotide sequence complementary to any of (a)-(e).

4. A vector comprising the nucleic acid molecule of claims 1, 2, or 3.

5. A host cell comprising the vector of claim 4.

6. The host cell of claim 5 that is a eukaryotic cell.

7. The host cell of claim 5 that is a prokaryotic cell.

8. A process of producing a huE3 α polypeptide comprising culturing the host cell of claim 5 under suitable conditions to express the polypeptide, and optionally isolating the polypeptide from the culture.

9. A polypeptide produced by the process of claim 8.

10. The process of claim 8, wherein the nucleic acid molecule comprises promoter DNA other than the promoter DNA for the native huE3 α polypeptide operatively linked to the DNA encoding the huE3 α polypeptide.

11. The isolated nucleic acid molecule according to claim 2 wherein the percent identity is determined using a computer program selected from the group consisting of GAP, BLASTP, BLASTN, FASTA, BLASTA, BLASTX, BestFit, and the Smith-Waterman algorithm.

12. A process for determining whether a compound inhibits huE3 α polypeptide activity or production comprising exposing a cell according to claims 5, 6, or 7 to the compound, and measuring huE3 α polypeptide activity or production in said cell.

13. An isolated polypeptide comprising the amino acid sequence set forth in SEQ ID NOS: 2 or 4.

14. An isolated polypeptide comprising the amino acid sequence selected from the group consisting of:

- (a) the mature amino acid sequence as set forth in SEQ ID NOS: 2 or 4 comprising a mature amino terminus at residue 1, optionally further comprising an amino-terminal methionine;
- (b) an amino acid sequence for an ortholog of SEQ ID NOS: 2 or 4;
- (c) an amino acid sequence that is at least about 70, 80, 85, 90, 95, 96, 97, 98, or 99 percent identical to the amino acid sequence of SEQ ID NOS: 2 or 4, wherein the polypeptide has an activity of the polypeptide set forth in SEQ ID NOS: 2 or 4;

(d) a fragment of the amino acid sequence set forth in SEQ ID NOS: 2 or 4 comprising at least about 25 amino acid residues, wherein the polypeptide has an activity of the polypeptide set forth in SEQ ID NOS: 2 or 4;

(e) an amino acid sequence for an allelic variant or splice variant of either the amino acid sequence as set forth in SEQ ID NOS: 2 or 4, or at least one of (a)-(c).

15. An isolated polypeptide of claim 14 wherein the amino acid sequence is a mouse ortholog set out in SEQ ID NO: 6.

16. An isolated polypeptide comprising the amino acid sequence selected from the group consisting of:

(a) the amino acid sequence as set forth in SEQ ID NOS: 2 or 4 with at least one conservative amino acid substitution, wherein the polypeptide has an activity of the polypeptide set forth in SEQ ID NOS: 2 or 4;

(b) the amino acid sequence as set forth in SEQ ID NOS: 2 or 4 with at least one amino acid insertion, wherein the polypeptide has an activity of the polypeptide set forth in SEQ ID NOS: 2 or 4;

(c) the amino acid sequence as set forth in SEQ ID NOS: 2 or 4 with at least one amino acid deletion, wherein the polypeptide has an activity of the polypeptide set forth in SEQ ID NOS: 2 or 4;

(d) the amino acid sequence as set forth in SEQ ID NOS: 2 or 4 which has a C- and/or N- terminal truncation, wherein the polypeptide has an activity of the polypeptide set forth in SEQ ID NOS: 2 or 4; and

(e) the amino acid sequence as set forth in SEQ ID NOS: 2 or 4, with at least one modification selected from the group consisting of amino acid substitutions, amino acid insertions, amino acid deletions, C-terminal truncation, and N-terminal truncation, wherein the polypeptide has an activity of the polypeptide set forth in SEQ ID NOS: 2 or 4.

17. An isolated polypeptide encoded by the nucleic acid molecule of claims 1, 2, or 3.

18. The isolated polypeptide according to claim 14 wherein the percent identity is determined using a computer program selected from the group consisting of GAP, BLASTP, BLASTN, FASTA, BLASTA, BLASTX, BestFit, and the Smith-Waterman algorithm.

5

19. An antibody produced by immunizing an animal with a peptide comprising an amino acid sequence of SEQ ID NOS: 2 or 4.

10

20. An antibody or fragment thereof that specifically binds the polypeptide of claims 13, 14, 15 or 16.

21. The antibody of claim 20 that is a monoclonal antibody.

15

22. A hybridoma that produces a monoclonal antibody that binds to a peptide comprising an amino acid sequence of SEQ ID NOS: 2 or 4.

23. A method of detecting or quantitating the amount of huE3 α polypeptide using the anti-huE3 α antibody or fragment of claims 18, 19, or 20.

20

24. A method for determining the presence and/or concentration of huE3 α polypeptide in a biological sample comprising the steps of:

(a) obtaining a biological sample;

(b) contacting the biological sample with an antibody according to claim 22 under conditions which allow for the antibody to bind to the huE3 α polypeptide;

25

(c) detecting antibody binding to huE3 α polypeptide in the biological sample, wherein the detection of antibody binding is indicative of the presence of huE3 α polypeptide in said biological sample; and

(d) comparing the level of antibody binding to huE3 α polypeptide within said biological sample and the level of the antibody binding to a known concentration of huE3 α polypeptide.

30

25. A selective binding agent or fragment thereof that specifically binds at least one polypeptide wherein said polypeptide comprises the amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence as set forth in SEQ ID NOS: 2 or 4; and
- (b) fragment of the amino acid sequence set forth in at least one of SEQ ID NOS: 2 or 4; or a naturally occurring variant thereof.

26. The selective binding agent of claim 25 that is an antibody or fragment thereof.

27. The selective binding agent of claim 25 that is a humanized antibody.

28. The selective binding agent of claim 25 that is a human antibody or fragment thereof.

29. The selective binding agent of claim 25 that is a polyclonal antibody or fragment thereof.

30. The selective binding agent claim 25 that is a monoclonal antibody or fragment thereof.

31. The selective binding agent of claim 25 that is a chimeric antibody or fragment thereof.

32. The selective binding agent of claim 25 that is a CDR-grafted antibody or fragment thereof.

33. The selective binding agent of claim 25 that is an antiidiotypic antibody or fragment thereof.

34. The selective binding agent of claim 25 which is a variable region fragment.

35. The variable region fragment of claim 25 which is a Fab or a Fab' fragment.

36. A selective binding agent or fragment thereof comprising at least one complementarity determining region with specificity for a polypeptide having the amino acid sequence of SEQ ID NOS: 2 or 4.

37. The selective binding agent of claim 25 which is bound to a detectable label.

38. A selective binding agent produced by immunizing an animal with a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NOS: 2 or 4.

39. A hybridoma that produces a selective binding agent capable of binding a polypeptide according to claims 1, 2, or 3.

40. A composition comprising the polypeptide of claims 13, 14, or 16 and a pharmaceutically acceptable formulation agent.

41. The composition of claim 40 wherein the pharmaceutically acceptable formulation agent is a carrier, adjuvant, solubilizer, stabilizer, or anti-oxidant.

42. The composition of claim 40 wherein the polypeptide comprises the mature amino acid sequence as set forth in SEQ ID NOS: 2 or 4.

43. A polypeptide comprising a derivative of the polypeptide of claims 13, 14, or 16.

44. The polypeptide of claim 43 which is covalently modified with a water-soluble polymer.

45. The polypeptide of claim 44 wherein the water-soluble polymer is selected from the group consisting of polyethylene glycol, monomethoxy-polyethylene glycol, dextran, cellulose, poly-(N-vinyl pyrrolidone) polyethylene glycol, propylene glycol homopolymers, polypropylene oxide/ethylene oxide co-polymers, polyoxyethylated polyols, and polyvinyl alcohol.

46. A composition comprising a nucleic acid molecule of claims 1, 2, or 3 and a pharmaceutically acceptable formulation agent.

47. A composition of claim 46 wherein said nucleic acid molecule is contained in a viral vector.

48. A viral vector comprising a nucleic acid molecule of claims 1, 2, or 3.

49. A fusion polypeptide comprising the polypeptide of claims 13, 14, or 16 fused to a heterologous amino acid sequence.

50. The fusion polypeptide of claim 49 wherein the heterologous amino acid sequence is an IgG constant domain or fragment thereof.

51. A method for treating, preventing or ameliorating a medical condition in a mammal resulting from decreased levels of huE3 α polypeptide comprising administering a huE3 α polypeptide to said mammal.

52. The method of claim 51 wherein the huE3 α polypeptide administered is the amino acid sequence set forth in SEQ ID NOS: 2 or 4 or a fragment thereof at least 25 amino acids or a homolog, analog or variant of said huE3 α polypeptide or fragment thereof.

53. The method of claim 51 wherein the huE3 α polypeptide administered has the amino acid set forth in SEQ ID NOS: 2 or 4 with at least one amino acid substitution.

54. The method of claim 51 wherein the huE3 α polypeptide administered has the amino acid sequence set forth in SEQ ID NOS: 2 or 4 with at least one amino acid deleted.

5 55. A method of diagnosing the existence or a susceptibility to a pathological condition in a subject caused by or resulting from abnormal levels of huE3 α in a mammalian subject comprising:

(a) determining the level of huE3 α in a biological, tissue or cellular sample; and

10 (b) comparing the level of huE3 α polypeptide in biological, tissue or cellular samples from normal subjects or the subject at an earlier time; wherein a susceptibility of an existence of a pathological condition is based on the presence or amount of expression of huE3 α polypeptide.

15 56. A device, comprising:

(a) a membrane suitable for implantation; and

(b) cells encapsulated within said membrane, wherein said cells secrete a protein of claims 13, 14, or 16; said membrane being permeable to said protein product and impermeable to materials detrimental to said cells.

20 57. A device, comprising:

(a) a membrane suitable for implantation; and

(b) The huE3 α polypeptide encapsulated within said membrane, wherein said membrane is permeable to the polypeptide.

25 58. A method of identifying a compound which binds to a polypeptide comprising:

(a) contacting the polypeptide of claims 13, 14, or 16 with a compound; and

30 (b) determining the extent of binding of the polypeptide to the compound.

59. A diagnostic reagent comprising a detectably labeled polynucleotide encoding the amino acid sequence set out in SEQ ID NOS: 2 or 4; or a fragment, variant or homolog thereof including allelic variants and spliced variants thereof.

5 60. The diagnostic reagent of claim 58, wherein said labeled polynucleotide is a first-strand cDNA.

61. A method for determine the presence of huE3 α nucleic acids in a biological sample comprising the steps of:

- 10 (a) providing a biological sample suspected of containing huE3 α nucleic acids;
- (b) contacting the biological sample with a diagnostic reagent according to claim 59 under conditions wherein the diagnostic reagent will hybridize with huE3 α nucleic acids contained in said biological sample;
- 15 (c) detecting hybridization between huE3 α nucleic acid in the biological sample and the diagnostic reagent; and
- (d) comparing the level of hybridization between the biological sample and diagnostic reagent with the level of hybridization between a known concentration of huE3 α nucleic acid and the diagnostic reagent.

20 62. A method for detecting the presence of huE3 α nucleic acids in a tissue or cellular sample comprising the steps of:

- (a) providing a tissue or cellular sample suspected of containing huE3 α nucleic acids;
- 25 (b) contacting the tissue or cellular sample with a diagnostic reagent according to claim 59 under conditions wherein the diagnostic reagent will hybridize with huE3 α nucleic acids;
- (c) detecting hybridization between huE3 α nucleic acid in the tissue or cellular sample and the diagnostic reagent; and
- 30 (d) comparing the level of hybridization between the tissue or cellular sample and diagnostic reagent with the level of hybridization between a known concentration of huE3 α nucleic acid and the diagnostic reagent.

63. The method of claim 59 wherein said polynucleotide molecule is DNA.

64. The method of claim 59 wherein said polynucleotide molecule is RNA.

65. A method of modulating levels of a polypeptide in an animal comprising administering to the animal the nucleic acid molecule of claims 1, 2, or 3.

66. A transgenic non-human mammal comprising the nucleic acid molecule of claims 1, 2, or 3.

ABSTRACT OF THE DISCLOSURE

The present invention relates to a novel polypeptide encoding a protein which is the full length human ortholog of E3 α ubiquitin ligase. The invention also relates to vector, host cells, antibodies and recombinant methods for producing the polypeptide. In addition, the invention discloses therapeutic, diagnostic and research utilities for these and related products.

5

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 1

SEQ ID NO:

6	mouse_E3αII	MASEMEPEVQ	AID-RSLL	EC	SAEEIAGRWL	QATDLNRE	VY	QHLAHCVPKI	49
4	human_E3αII	MASELEPEVQ	AID-RSLL	EC	SAEEIAGKWL	QATDLTRE	VY	QHLAHYVPKI	49
15	mouse_E3αI	MADEEMDGAE	RMDVSPEP	PL	APQRPASW	WD	QQVDFYTA	FL	50
2	human_E3αI	MADEEAGGTE	RMEISAE	LPQ	TPQRLASW	WD	QQVDFYTA	FL	50
	Consensus	MA.E.....	..D....L..	A..W.		Q..D.....	.HLA..VP.I	50
6	mouse_E3αII	YCRGPNPFPQ	KEDTLAQHIL		LGPMEWYICA	EDPALGF	PKL	EQANKPSHLC	99
4	human_E3αII	YCRGPNPFPQ	KEDMLAQHVL		LGPMEWYLCG	EDPAFGF	PKL	EQANKPSHLC	99
15	mouse_E3αI	YFAEMDPDLE	KQEESVQMSI		LTPLEWYLF	G	EDPDICLE	KL	99
2	human_E3αI	YFAEMDPDLE	KQEESVQMSI		FTPLEWYLF	G	EDPDICLE	KL	99
	Consensus	Y.....P...	K.....Q...		L.P.EWYL.G		EDP.....KLLC	100
6	mouse_E3αII	GRVFKVGEPT	YSCRDCAVDP		TCVLCMECF	L	GSIH	R	149
4	human_E3αII	GRVFKVGEPT	YSCRDCAVDP		TCVLCMECF	L	GSIH	R	149
15	mouse_E3αI	GKVF	KSGETT	YSCRDC	AIDP	TCVLCMD	CFQ	SSVHKNHRYK	149
2	human_E3αI	GRVFKSGETT	YSCRDC	AIDP	TCVLCMD	CFQ	DSVHKNHRYK	MHTSTGGGFC	149
	Consensus	GRVFK.GE.T	YSCRDC	.DP	TCVLCM	.CF.	.S.H..HRY.	M.TS.GGGFC	150
6	mouse_E3αII	DCGDTEAWKE	GPYCQKHKL	S	SSEVVEEEDP	LVHLS	EDVIA	RTYNIFAIMF	199
4	human_E3αII	DCGDTEAWKE	GPYCQKHEL	N	TSEIEEEDP	LVHLS	EDVIA	RTYNIFAITF	199
15	mouse_E3αI	DCGDTEAWKT	GPFCVDHEPG		RAGTTKESLH	-CPLNEE	VIA	QARRIFPSVI	198
2	human_E3αI	DCGDTEAWKT	GPFCVNHEPG		RAGTIKENS	R	-CPLNEEVIV	QARKIFPSVI	198
	Consensus	DCGDTEAWK.	GP.C..HE..	E...		...L.E.VIAIF....	200
6	mouse_E3αII	RYAVDILTWE	KESELPEDLE		VAEKSDTYYC	MLFNDE	EVHTY	EQVIYTLQKA	249
4	human_E3αII	RYAVEILTWE	KESELPADLE		MVEKSDTYYC	MLFNDE	EVHTY	EQVIYTLQKA	249
15	mouse_E3αI	KYIVEMTIWE	EEKELPPELQ		IREKNERYYC	VLFNDE	EHHSY	DHVIYSLQRA	248
2	human_E3αI	KYVEMTIWE	EEKELPPELQ		IREKNERYYC	VLFNDE	EHHSY	DHVIYSLQRA	248
	Consensus	.Y.VE...WE	.E.ELP..L.		..EK...YYC	.LFNDE.H.Y	.VIY.LQ.A		250
6	mouse_E3αII	VNCTQKEAIG	FATTVD	R	DGR	RPVRYG	D	FQY	298
4	human_E3αII	VNCTQKEAIG	FATTVD	R	DGR	RSVRYG	D	FQY	298
15	mouse_E3αI	LDCELA	EAQ	L	HTTAIDKEGR	RAVKAG	VYAT	CQEAKEDIKS	298
2	human_E3αI	LDCELA	EAQ	L	HTTAIDKEGR	RAVKAG	AYAA	CQEAKEDIKS	298
	Consensus	..C...EA..	..T..D..GR		R.V..G....	C..AK..I..PL		300
6	mouse_E3αII	KVQVMHSSVA	AHQNFGLKAL		SWLG	SVIGYS	DGLRRILCQV	GLQEGPDGEN	348
4	human_E3αII	KVQVMHSSIV	AHQNFGLKLL		SWLGS	IIGYS	DGLRRILCQV	GLQEGPDGEN	348
15	mouse_E3αI	HVEVLH	SVVM	AHQKFALRLG		SWMNKIMS	YS	SDFRQIFCQA	348
2	human_E3αI	HVEVLHSEIM	AHQKFALRLG		SWMNKIMS	YS	SDFRQIFCQA	CLREEPDSEN	348
	Consensus	.V.V.HS...	AHQ.F.L.L.		SW...I..YS		...R.I.CQ.	.L.E.PD.EN	350
6	mouse_E3αII	SSLVDRLMLN	DSKLWKGARS		VYHQLFMSSL	LMDLKYKKLF	ALRFAKNYRQ		398
4	human_E3αII	SSLVDRLMLS	DSKLWKGARS		VYHQLFMSSL	LMDLKYKKLF	AVRFAKNYQQ		398
15	mouse_E3αI	PCLISRLMLW	DAKLYKGARK		ILHELIFSSF	FME	MEYKKLF	AMEFVKYYKQ	398
2	human_E3αI	PCLISRLMLW	DAKLYKGARK		ILHELIFSSF	FME	MEYKKLF	AMEFVKYYKQ	398
	Consensus	..L..RLML.	D.KL.KGAR.		..H.L..SS.	.M...YKKLF	A..F.K.Y.Q		400
6	mouse_E3αII	LQRDFMEDDH	ERAVSVTALS		VQFFTAPT	LA	RMLL	TEENLM	448
4	human_E3αII	LQRDFMEDDH	ERAVSVTALS		VQFFTAPT	LA	RML	TEENLM	448
15	mouse_E3αI	LQKEYISDDH	ERSISITALS		VQMLTVPT	LA	RHLIEEQNVI	SVITETLLEV	448
2	human_E3αI	LQKEYISDDH	DRSISITALS		VQMFTVPT	LA	RHLIEEQNVI	SVITETLLEV	448
	Consensus	LQ.....DDH	ER..S.TALS		VQ.FT.PTLA		R.LI.E.N..	SVI..T...	450

Figure 1 (continued)

SEQ ID NO:

6	mouse_E3αII	LKHRDAQGRF	QFERYTALQA	FKFRRVQSLI	LDLKYVLISK	PTIEWSDELRO	498
4	human_E3αII	LRHRDAQGRF	QFERYTALQA	FKFRRVQSLI	LDLKYVLISK	PTIEWSDELRO	498
15	mouse_E3αI	LPEYLDNRN-	KFN-FQGYSQ	DKLGRVYAVI	CDLKYILISK	PVIWTERLRA	496
2	human_E3αI	LPEYLDNRN-	KFN-FQGYSQ	DKLGRVYAVI	CDLKYILISK	PTIWTERLRM	496
	Consensus	L.....	.F.....	.K..RV...I	.DLKY.LISK	PT.W...LR.	500
6	mouse_E3αII	KFLQGFDAFL	ELLKCMQGM	PITRQVGQHI	EMEPWEAAAF	TLQMKLTHVI	548
4	human_E3αII	KFLEGFDAFL	ELLKCMQGM	PITRQVGQHI	EMEPWEAAAF	TLQMKLTHVI	548
15	mouse_E3αI	QFLEGFRSFL	KILTCMQGME	EIRRVQVGQHI	EVDPDWEAAI	AIQMQLKNIL	546
2	human_E3αI	QFLEGFRSFL	KILTCMQGME	EIRRVQVGQHI	EVDPDWEAAI	AIQMQLKNIL	546
	Consensus	.FLEGF..FL	..L.CMQGM.	.I.RQVGQHI	E..P.WEAA.	..QM.L....	550
6	mouse_E3αII	SMVQDWCALD	EKVLIEAYKK	CLAVLTQCHG	GFTDGEQPIT	LSICGHSVET	598
4	human_E3αII	SMMQDWCASD	EKVLIEAYKK	CLAVLMQCHG	GYTDGEQPIT	LSICGHSVET	598
15	mouse_E3αI	LMFQEWCACD	EDLLLVAIKE	CHKAVMRCST	NFMSSTKTIV	VQLCGHSLET	595
2	human_E3αI	LMFQEWCACD	EDLLLVAIKE	CHKAVMRCST	SFISSTKTIV	VQSCGHSLET	595
	Consensus	:M.Q.WCA.D	E..L..AYK.	C....M.C..	.F.....	...CGHS.ET	600
6	mouse_E3αII	IRYCVSQEKV	SIHLPLSRLL	AGLHVLLSKS	EVAYKFPPELL	PLSELSPMML	648
4	human_E3αII	IYCVSQEKV	SIHLPVSRLL	AGLHVLLSKS	EVAYKFPPELL	PLSELSPMML	648
15	mouse_E3αI	KSYKVSEDLV	SIHLPLSRTL	AGLHVRLSRL	GAISRLHEFV	PFDSFQVEVL	645
2	human_E3αI	KSYRVSEDLV	SIHLPLSRTL	AGLHVRLSRL	GAVSRLHEFV	SFEDFQVEVL	645
	Consensus	..Y.VS...V	SIHLP.SR.L	AGLHV.LS..E..	P.....L	650
6	mouse_E3αII	IEHPLRCLVL	CAQVHACMWR	RNGFSLVNQI	YYYHNVKCR	EMFDKDIVML	698
4	human_E3αII	IEHPLRCLVL	CAQVHACMWR	RNGFSLVNQI	YYYHNVKCR	EMFDKDVVML	698
15	mouse_E3αI	VEYPLRCLVL	VAQVVAEMWR	RNGLSLISQV	FYYQDVKCRE	EMYDKDIIML	695
2	human_E3αI	VEYPLRCLVL	VAQVVAEMWR	RNGLSLISQV	FYYQDVKCRE	EMYDKDIIML	695
	Consensus	.E.PLRCLVL	.AQV.A.MWR	RNG.SL..Q.	.YY..VKCR.	EM.DKDI.ML	700
6	mouse_E3αII	QTGVSMMDPN	HFLMIMLSRF	ELYQLFSTPD	YGKRFSSEVT	HKDQVQNNNT	748
4	human_E3αII	QTGVSMMDPN	HFLMIMLSRF	ELYQIFSTPD	YGKRFSSEIT	HKDQVQNNNT	748
15	mouse_E3αI	QIGASIMDPN	KFLLLVLQRY	EL-----TDA	FNKTIST--K	DQDLIKQYNT	738
2	human_E3αI	QIGASIMDPN	KFLLLVLQRY	EL-----AEA	FNKTIST--K	DQDLIKQYNT	738
	Consensus	Q.G.S.MDPN	.FL...L.R.	EL.....T..	..K..S....	..D...Q.NT	750
6	mouse_E3αII	LIEEMLYLII	MLVGERFNP	VGQVAATDEI	KREIIHQLSI	KPMAHSELVK	798
4	human_E3αII	LIEEMLYLII	MLVGERFSPG	VGQVNATDEI	KREIIHQLSI	KPMAHSELVK	798
15	mouse_E3αI	LIEEMLQVLI	YIVGERYVPG	VGNVTREEVI	MREITHLLCI	EPMPHSAIAR	788
2	human_E3αI	LIEEMLQVLI	YIVGERYVPG	VGNVTKEEVT	MREIIHLLCI	EPMPHSAIAK	788
	Consensus	LIEEML...I	..VGER..PG	VG.V.....I	.REIIH.L.I	.PM.HS...K	800
6	mouse_E3αII	SLPEDENKET	GMESVIESVA	HFKKPGLTGR	GMVELKPECA	KEFNLYFYHF	848
4	human_E3αII	SLPEDENKET	GMESVIEAVA	HFKKPGLTGR	GMVELKPECA	KEFNLYFYHF	848
15	mouse_E3αI	NLPENENNET	GLENVINKVA	TFKKPGVSGH	GVYELKDESL	KDFNMYFYHY	838
2	human_E3αI	NLPENENNET	GLENVINKVA	TFKKPGVSGH	GVYELKDESL	KDFNMYFYHY	838
	Consensus	.LPE.EN.ET	G.E.VI..VA	.FKKPG..G.	G.YELK.E..	K.FN.YFYH.	850
6	mouse_E3αII	SRAEQSKAEE	AQRKLKRENK	EDTALPPPAP	PPFCPLFASL	VNIIQCDVML	898
4	human_E3αII	SRAEQSKAEE	AQRKLKRONR	EDTALPPPVL	PPFCPLFASL	VNIIQSDVML	898
15	mouse_E3αI	SKTQHSKAEH	MQKKRRKQEN	KDEALPPPPP	PEFCPAFSKV	VNLLSCDVMI	888
2	human_E3αI	SKTQHSKAEH	MQKKRRKQEN	KDEALPPPPP	PEFCPAFSKV	INLLNCDIMM	888
	Consensus	S....SKAE.	.Q.K...Q..	.D.ALPPP..	P.FCP.F...	VN.L.CDVM.	900

Figure 1 (continued)

SEQ ID NO:

6	mouse_E3αII	YIMGTLQWA	VEHHGSAWSE	SMLQRVLHLI	GMALQEEKHH	LENAVEGHVQ	948
4	human_E3αII	CIMGTLQWA	VEHNGYAWSE	SMLQRVLHLI	GMALQEEKQH	LENVTEEHVV	948
15	mouse_E3αI	YILRTIFERA	VDTESNLWTE	GMLQMAFHIL	ALGLLEEKQQ	LQKAPEEEV-	937
2	human_E3αI	YILRTVFERA	IDTDSNLWTE	GMLQMAFHIL	ALGLLEEKQQ	LQKAPEEEV-	937
	Consensus	YI..TI...A	V.....W.E	.MLQ...H..	...L.EEKQ.	L..A.EE.V.	950
6	mouse_E3αII	TFTFTQKISK	PGDAPHNSPS	ILAMLETLQN	APSLEAHKDM	IRWLLKMFNA	998
4	human_E3αII	TFTFTQKISK	PGEAPKNSPS	ILAMLETLQN	APYLEVHKDM	IRWILKTFNA	998
15	mouse_E3αI	AFDFYHKASR	LGSSAMNAQN	IQMLLERLKG	IPQLEGQKDM	ITWILQMFDT	987
2	human_E3αI	TFDFYHKASR	LGSSAMNIQM	L---LEKLKG	IPQLEGQKDM	ITWILQMFDT	984
	Consensus	TF.F..K.S.	.G....N...	I...LE.L..	.P.LE..KDM	I.WIL.MF..	1000
6	mouse_E3αII	IKKIRE--CS	SSSPVAEAE	TIMEESSRDK	DKAERKRKAE	IARLRREKIM	1046
4	human_E3αI	VKKMRE--SS	PTSPVAETEG	TIMEESSRDK	DKAERKRKAE	IARLRREKIM	1046
15	mouse_E3αI	VKRLREKSCL	VVATTSGLEC	IKSEEITHDK	EKAERKRKAE	AARLHRQKIM	1037
2	human_E3αI	VKRLREKSCL	IVATTSGSES	IKNDEITHDK	EKAERKRKAE	AARLHRQKIM	1034
	Consensus	VK..RE..C.E.	...EE...DK	.KAERKRKAE	.ARL.R.KIM	1050
6	mouse_E3αII	AQMSEMQRHF	IDENKELFQQ	TLELDTSASA	TL--DSSPPV	SDAALTALGP	1094
4	human_E3αII	AQMSEMQRHF	IDENKELFQQ	TLELDASTSA	VL--DHSPVA	SDMTLTALGP	1094
15	mouse_E3αI	AQMSALQKNF	IETHKLMYDN	TSEVTGKEDS	IMEEESTSAV	SEASRIALGP	1087
2	human_E3αI	AQMSALQKNF	IETHKLMYDN	TSEMPGKEDS	IMEEESTPAV	SDYSRIALGP	1084
	Consensus	AQMS..Q..F	I...K.....	T.E.....S.P.V	SD....ALGP	1100
6	mouse_E3αII	AQTQVPEPRQ	FVTCILCQEE	QEVTVGSRAM	VLAAFVQRST	VLSKDRTKTI	1144
4	human_E3αII	TQTQVPEQRQ	FVTCILCQEE	QEVKVESRAM	VLAAFVQRST	VLSKNRSKFI	1144
15	mouse_E3αI	KRGPAVTEKE	VLTCILCQEE	QEVKLENNAM	VLSACVQKST	ALTQHRGKPV	1137
2	human_E3α	KRGPSVTEKE	VLTCILCQEE	QEVKIENNAM	VLSACVQKST	ALTQHRGKPI	1134
	ConsensusTCILCQEE	QEVK.E..AM	VL.A.VQ.ST	.L...R.K.I	1150
6	mouse_E3αII	AD-PEKYDPL	FMHPDLSCGT	HTGSCGHVMH	AHCWQRYFDS	VQAKEQRRQQ	1193
4	human_E3αII	QD-PEKYDPL	FMHPDLSCGT	HTSSCGHIMH	AHCWQRYFDS	VQAKEQRRQQ	1193
15	mouse_E3αI	DHLGETLDPL	FMDPDLAGHT	YTGSCGHVMH	AVCWQKYFEA	VQ---LSSQQ	1184
2	human_E3αI	ELSGEALDPL	FMDPDLAYGT	YTGSCGHVMH	AVCWQKYFEA	VQ---LSSQQ	1181
	ConsensusE..DPL	FM.PDL..GT	.TGSCGHVMH	A.CWQ.YF..	VQ.....QQ	1200
6	mouse_E3αII	RLRLHTSYDV	ENGEFLCPLC	ECLSNTVIPL	L-LPPRSILS	RRLN-FSDQP	1241
4	human_E3αII	RLRLHTSYDV	ENGEFLCPLC	ECLSNTVIPL	L-LPPRNIFN	NRLN-FSDQP	1241
15	mouse_E3αI	RIHVDL-FDL	ESGEYLCPLC	KSLCNTVIPI	IPLQPQKINS	ENAEALAQLL	1233
2	human_E3αI	RIHVDL-FDL	ESGEYLCPLC	KSLCNTVIPI	IPLQPQKINS	ENADALAQLL	1230
	Consensus	R.....D.	E.GE.LCPLC	..L.NTVIP.	..L.P..I.S	1250
6	mouse_E3αII	DLAQWTRAVT	QQIKVVQMLR	RKHNAA-DTS	SSSEDTEAMNI	IPIPEGFRPD	1290
4	human_E3αII	NLTQWIRTIS	QQIKALQFLR	KEESTP-NNA	STKNSENVDE	LQLPEGFRPD	1290
15	mouse_E3αI	TLARWIQTVL	ARISGYNIKH	AKGEAPAVPV	LFNQMGDST	FEFHSILSFG	1283
2	human_E3αI	TLARWIQTVL	ARISGYNIRH	AKGENP-IP	FFNQMGDST	LEFHSILSFG	1279
	Consensus	.LA.WI.TV.	..I.....	.K...P-...	1300
6	mouse_E3αII	FYPRNPYSDS	IKEMLTTFGT	AAYKVGKLVH	PNEGDPVRPI	LCWGTCAYTI	1340
4	human_E3αII	FRPKIPYSES	IKEMLTTFGT	ATYKVGKLVH	PNEEDPRVPI	MCWGSCAYTI	1340
15	mouse_E3αI	VQSSVKYSNS	IKEMVILFAT	TIYRIGLKVP	PDELDPRVPM	MTWSTCAFTI	1333
2	human_E3αI	VESSIKYSNS	IKEMVILFAT	TIYRIGLKVP	PDERDPRVPM	LTWSTCAFTI	1329
	ConsensusYS.S	IKEM...F.T	..Y..GLKV.	P.E.DPRVP.	..W.TCA.TI	1350

Figure 1 (continued)

SEQ ID NO:

6	mouse_E3aII	QSIERILSDE	EKPVFGLPLPC	RLDDCLRSLT	RFAAAHWTVA	LLPVVQGHFC	1390
4	human_E3aII	QSIERILSDE	DKPLFGPLPC	RLDDCLRSLT	RFAAAHWTVA	SVSVVQGHFC	1390
15	mouse_E3aI	QAIENLLGDE	GKPLFGALQN	RQHSGLKALM	QFAVAQRATC	PQVLIHKHLA	1383
2	human_E3aI	QAIENLLGDE	GKPLFGALQN	RQHNGLKALM	QFAVAQRITC	PQVLIQKHLV	1379
	Consensus	Q.IE..L.DE	.KPLFG.L..	R....L..L.	.FA.A.....Q.H..	1400
6	mouse_E3aII	KLFASLVPSD	SYEDLPCILD	IDMFHLLVGL	VLAFPALQCQ	D---FSGSSL	1437
4	human_E3aII	KLFASLVPSD	SHEELPCILD	IDMFHLLVGL	VLAFPALQCQ	D---FSGISL	1437
15	mouse_E3aI	RLLSVILPNL	QSENTPGLLS	VDLFHVLVGA	VLAFPSLYWD	DTVDLQPSPL	1433
2	human_E3aI	RLLSVVLPNI	KSEDTPCLLS	IDLFHVLVGA	VLAFPSLYWD	DPVDLQPSV	1429
	Consensus	.L.....PN.	..E..PC.L.	ID.FH.LVG.	VLAFP.L...	D.....SSL	1450
6	mouse_E3aII	ATG--DLHIF	HLVTMAHIVQ	ILLTSCTEEN	---GMDQENP	TGEEELAILS	1482
4	human_E3aII	GTG--DLHIF	HLVTMAHIIQ	ILLTSCTEEN	---GMDQENP	PCEEESAVLA	1482
15	mouse_E3aI	SSSYNHLYLF	HLITMAHMLQ	ILLTTDTDLS	PGPPLAEGEE	DSEEARCASA	1483
2	human_E3aI	SSSYNHLYLF	HLITMAHMLQ	ILLTVDTGL-	---PLAQVQE	DSEEAHSAAS	1475
	ConsensusL..F	HL.TMAH..Q	ILLT..T...	-----Q...	..EE.....	1500
6	mouse_E3aII	LHKTLLHQYT	SALKEAPSGW	HLWRSVRAAI	MPFLKCSALF	FHYLNGVPAP	1532
4	human_E3aII	LYKTLLHQYT	SALKEIPSGW	HLWRSVRAGI	MPFLKCSALF	FHYLNGVPSP	1532
15	mouse_E3aI	FFVEVSQHTD	GLTGCGAPGW	YLWLSLRNGI	TPYLRC AALL	FHYLLGVAPP	1533
2	human_E3aI	FFAEISQYTS	GSIGCDIPGW	YLWVSLKNGI	TPYLRC AALF	FHYLLGVTPP	1525
	ConsensusQYT.GW	.LW.S.R.GI	.P.L.C.ALF	FHYL.GV..P	1550
6	mouse_E3aII	PDLQV-SGTS	HFEHLCNYLS	LPTNLIHLFQ	ENSDIMNSLI	ESWCQNSEVK	1581
4	human_E3aII	PDIQV-PGTS	HFEHLC SYLS	LPNNLICLFQ	ENSEIMNSLI	ESWCRNSEVK	1581
15	mouse_E3aI	EELFANSAEG	EFSALCSYLS	LPTNFLLLFQ	EYWDTIRPLL	QRWCQDPALL	1583
2	human_E3aI	EELHTNSAEG	EYSALCSYLS	LPTNFLLLFQ	EYWDTVRPLL	QRWCADPALL	1575
	Consensus	..L...S...	.F..LCSYLS	LPTNL..LFQ	E..D....L.	..WC.....	1600
6	mouse_E3aII	RYLNGERGAI	SYPRGANKLI	DLPEDYSSLI	NQASNFSCPK	SGGDKSRAPT	1631
4	human_E3aII	RYLEGERDAI	RYPRESNKLI	NLPEDYSSLI	NQASNFSCPK	SGGDKSRAPT	1631
15	mouse_E3aI	KSLKQKSAVV	RYPRKRNSLI	ELPEDYSCLL	NQASHFRCPR	SADDERKHPV	1633
2	human_E3aI	NCLKQKNTVV	RYPRKRNSLI	ELPDDYSCLL	NQASHFRCPR	SADDERKHPV	1625
	Consensus	..L.....	RYPR..N.LI	.LPEDYS.L.	NQAS.F.CP.	S..D....P.	1650
6	mouse_E3aII	LCLVCGSLLC	SQSYCCQABL	BGEDVGACTA	HTYSCGSGAG	IFLVRRECQV	1681
4	human_E3aII	LCLVCGSLLC	SQSYCCQTEL	BGEDVGACTA	HTYSCGSGVG	IFLVRRECQV	1681
15	mouse_E3aI	LCLFCGAILC	SONICCQEIV	NGEEVGACVF	HALHCGAGVC	IFLKIRECRV	1683
2	human_E3aI	LCLFCGAILC	SONICCQEIV	NGEEVGACIF	HALHCGAGVC	IFLKIRECRV	1675
	Consensus	LCL.CG..LC	SQ..CCQ...	.GE.VGAC..	H...CG.GV.	IFL..REC.V	1700
6	mouse_E3aII	LFLAGKTKGC	FYSPPYLDDY	GETDQGLRRG	NPLHLCQERF	RKIQKLWQQH	1731
4	human_E3aII	LFLAGKTKGC	FYSPPYLDDY	GETDQGLRRG	NPLHLCKERF	KKIQKLWHQH	1731
15	mouse_E3aI	VLVEGKARGC	AYPAPYLDEY	GETDPGLKRG	NPLHLSRERY	RKLHLVWQQH	1733
2	human_E3aI	VLVEGKARGC	AYPAPYLDEY	GETDPGLKRG	NPLHLSRERY	RKLHLVWQQH	1725
	ConsensusGK..GC	.Y..PYLD.Y	GETD.GL.RG	NPLHL..ER.	RK....WQQH	1750
6	mouse_E3aII	SITEEIGHAQ	EANQTLVGID	WQHL			1755
4	human_E3aII	SVTEEIGHAQ	EANQTLVGID	WQHL			1755
15	mouse_E3aI	CIIEEIARSQ	ETNQMLFGFN	WQLL			1757
2	human_E3aI	CIIEEIARSQ	ETNQMLFGFN	WQLL			1749
	Consensus	.I.EEI...Q	E.NQ.L.G..	WQ.L			1774

Figure 2
The Expression Profile of huE3 α -II in Human Tissues

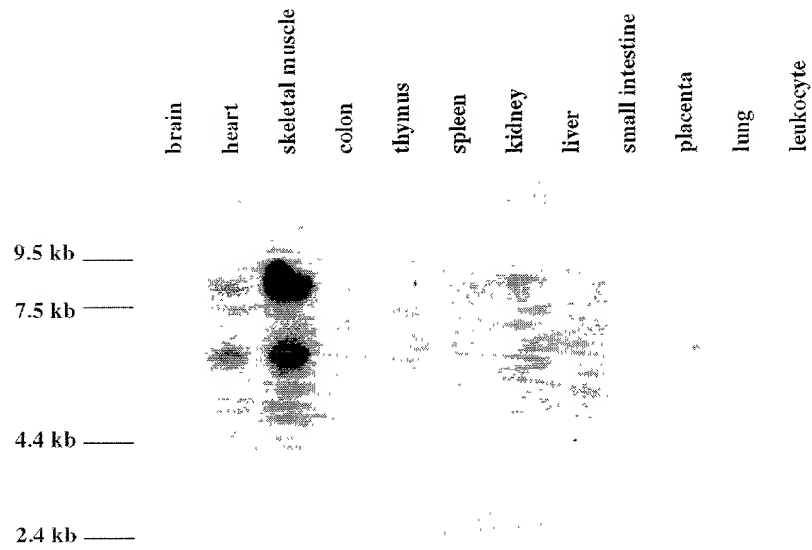


Figure 3

The Expression Profile of huE3 α -I in Human Tissues

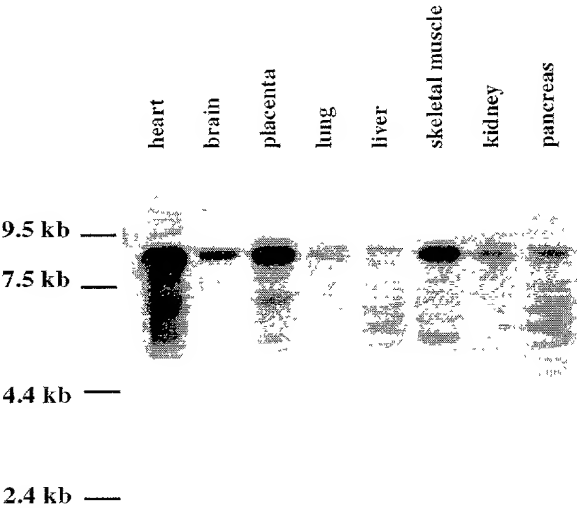


Figure 4- Ubiquitination of Endogenous Proteins

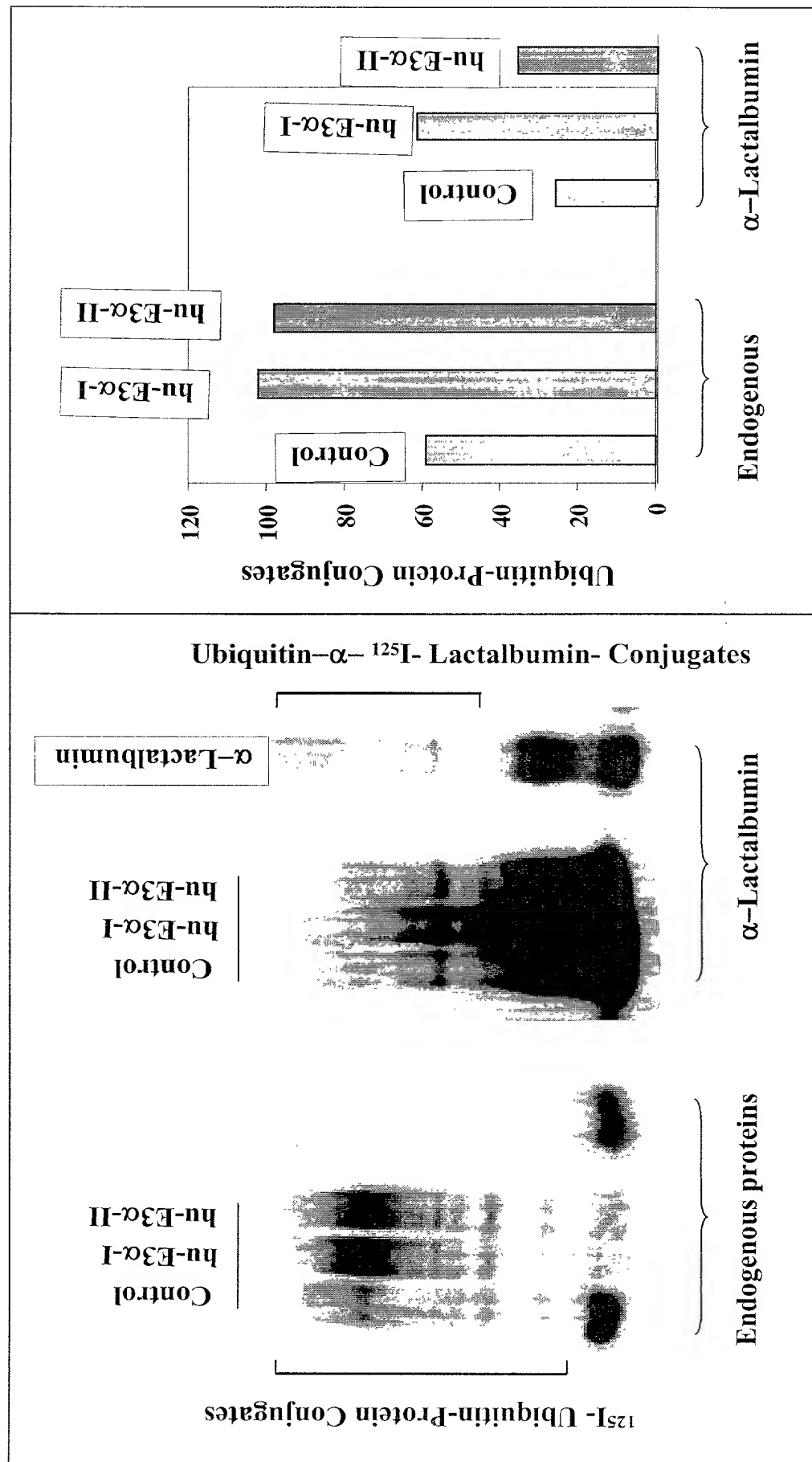


Figure 4- Ubiquitination of Endogenous Proteins

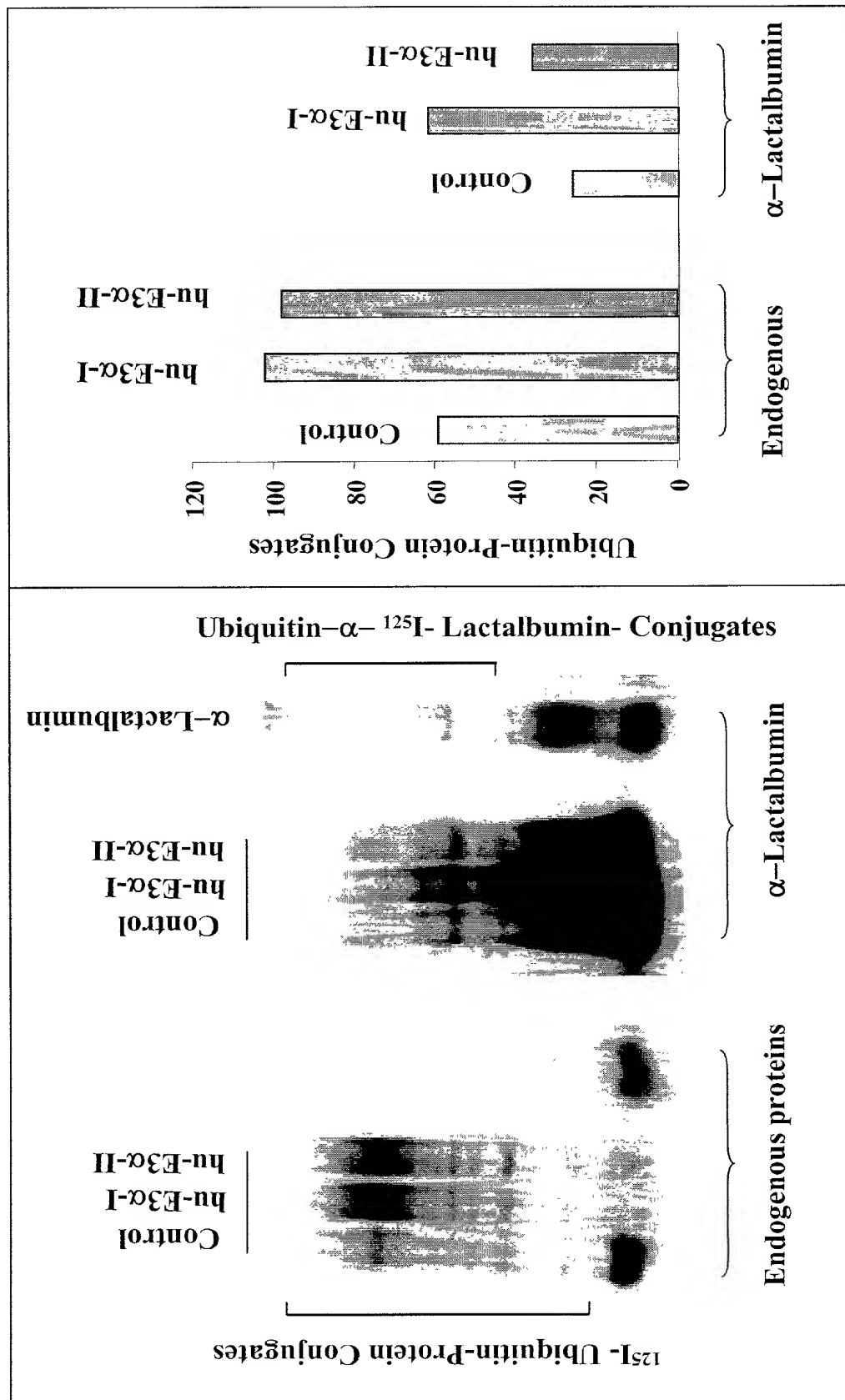


Figure 5

Transfection of Human E3a-I or E3a-II cDNA Stimulates Ubiquitin Conjugation in Cultured Muscle Cell Lines

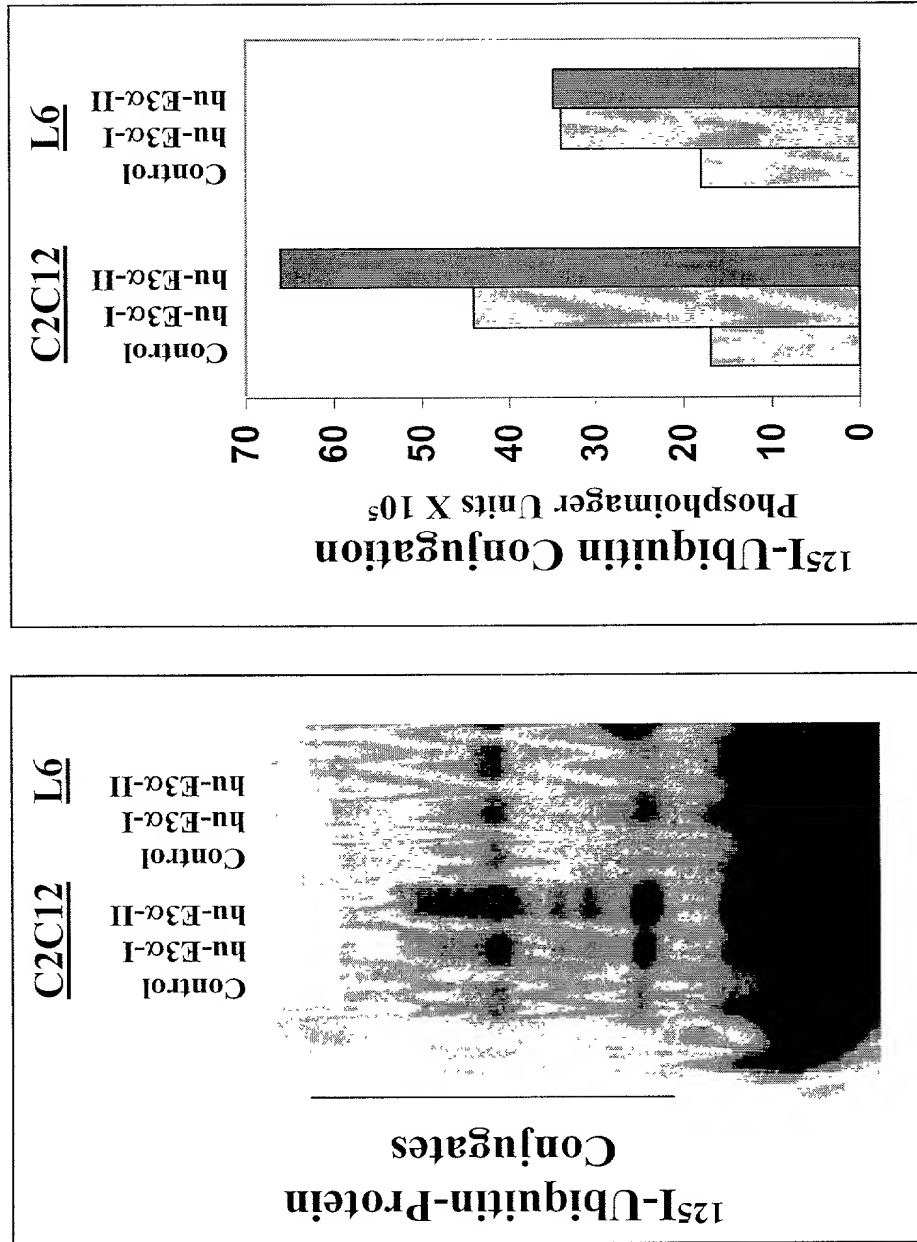


Figure 6 ^{125}I -Ubiquitin Conjugation to Muscle Proteins and Its Sensitivity to E3 α Inhibitor in Skeletal Muscle Extracts

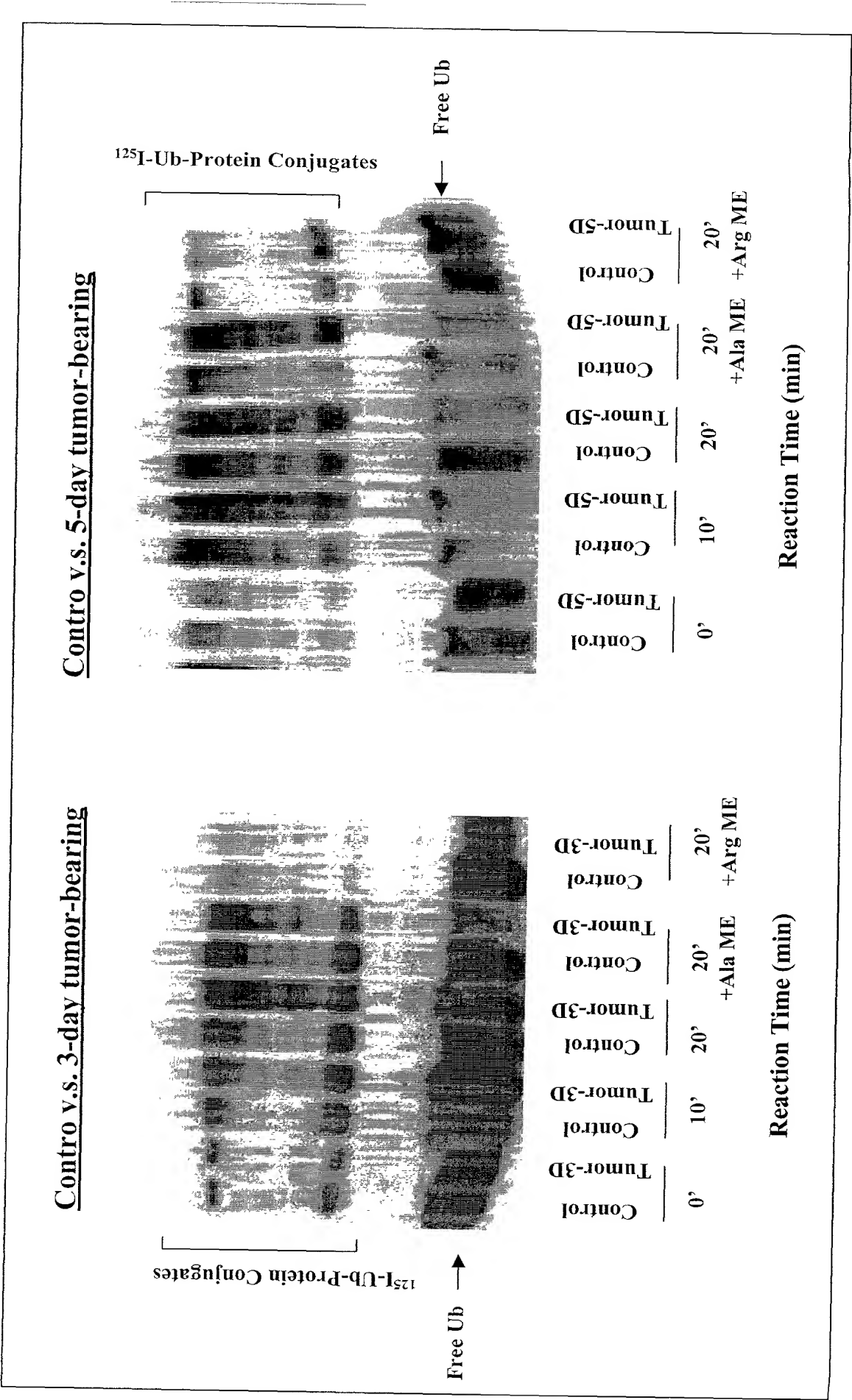


Figure 7

Rates of Ubiquitination of N-end Rule Substrate α -Lactalbumin in Skeletal Muscle Extracts

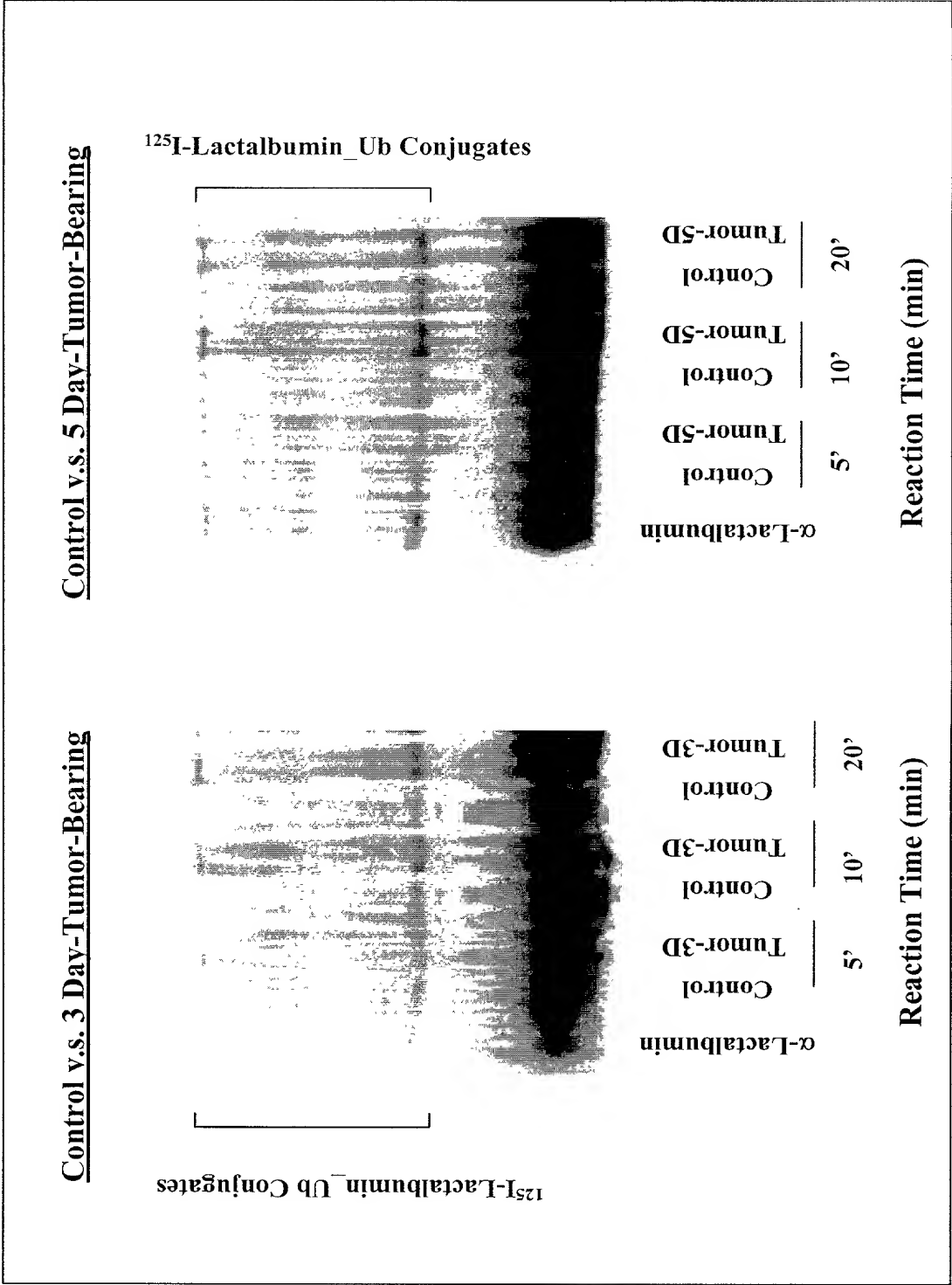


Figure 8

**Northern blot analysis of E3 α -I & E3 α -II expression
in gastrocnemius muscles in YAH-130 experimental cachexia model**

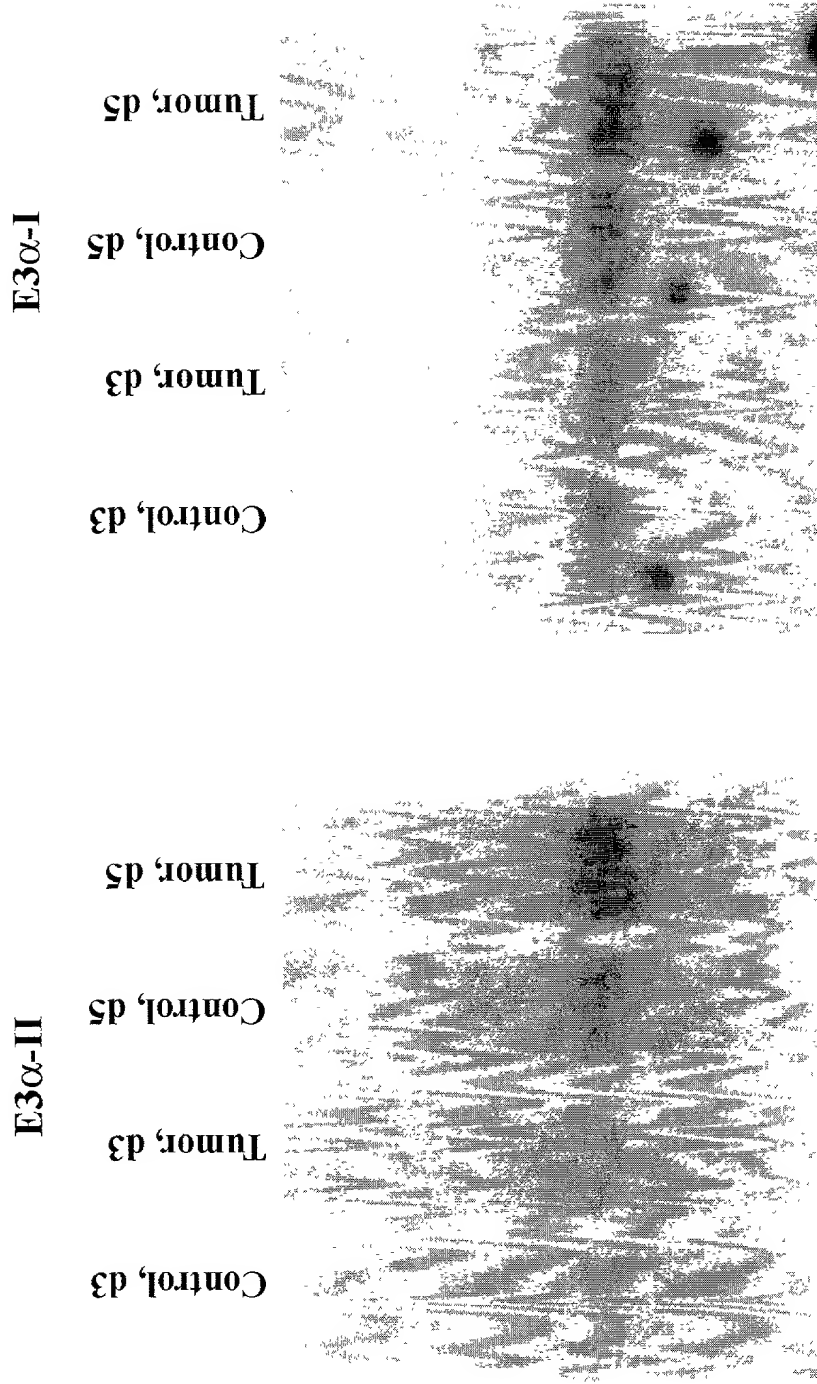


Figure 9

Northern blot analysis of E3 α -I and E3 α -II expression in gastrocnemius muscle and cardiac muscle in C26 experimental cachexia model

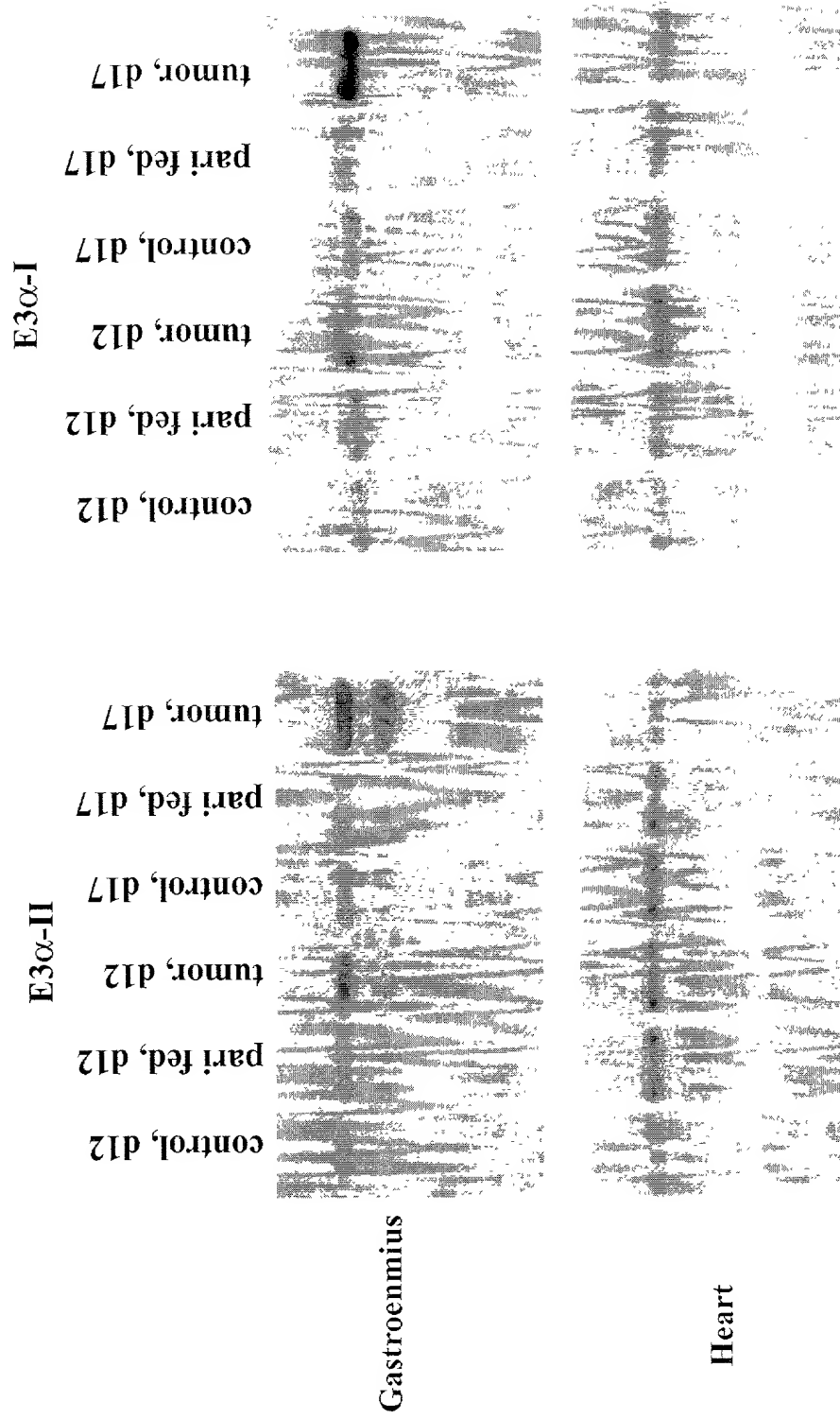


Figure 10

Proinflammatory cytokines TNF- α and IL-6
induce E3 α -II Expression in C2C12 myotube culture

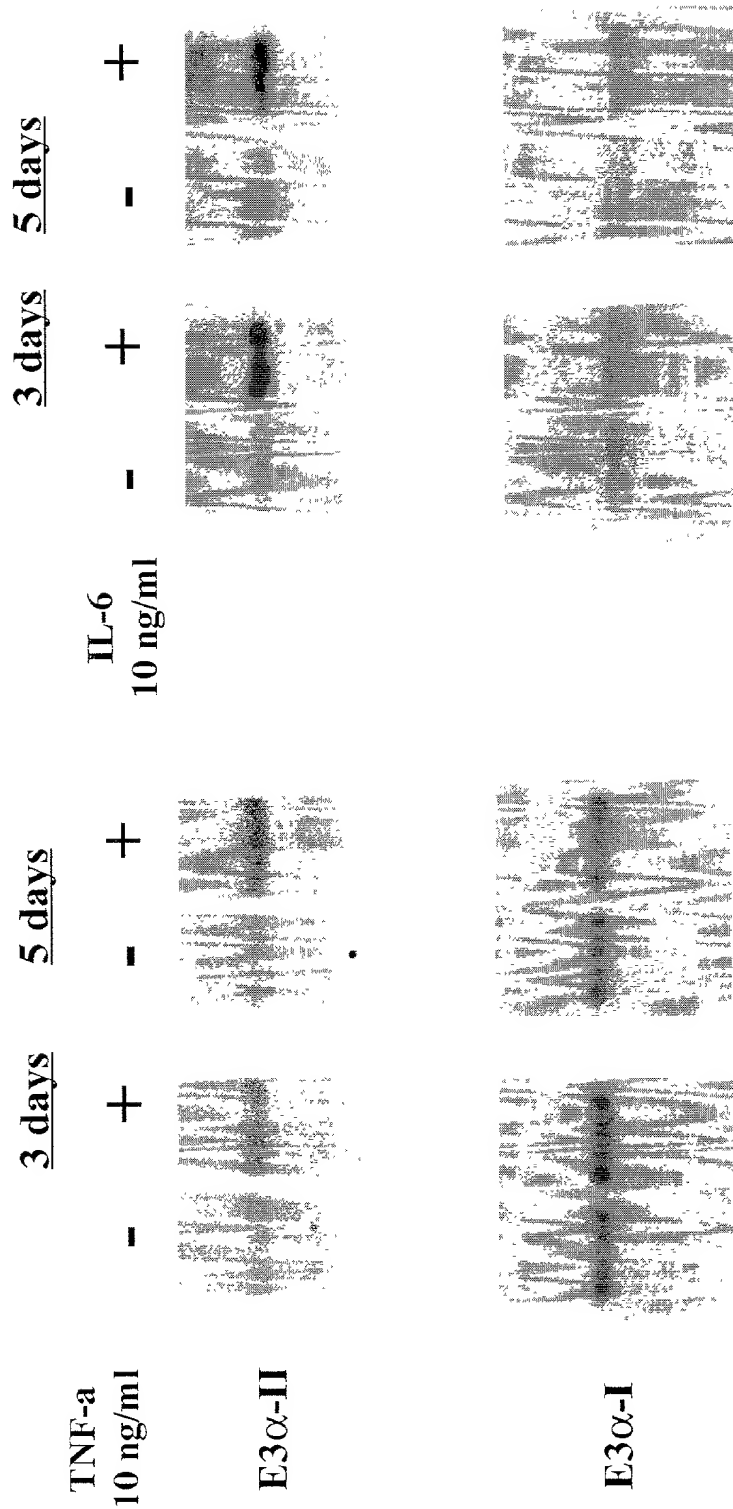


Figure 11

IL-6 Elicits Accelerated Ubiquitination in C2C12 Myotube Cultures

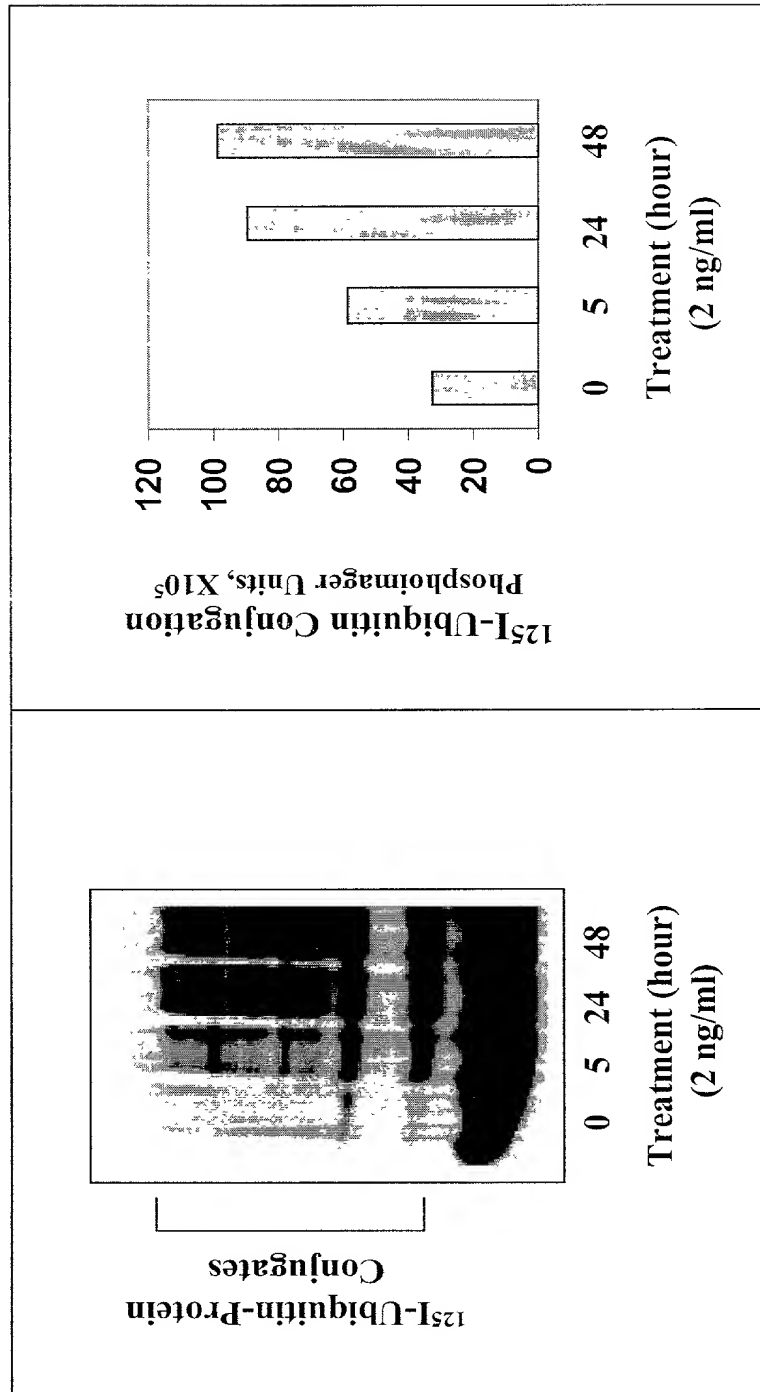
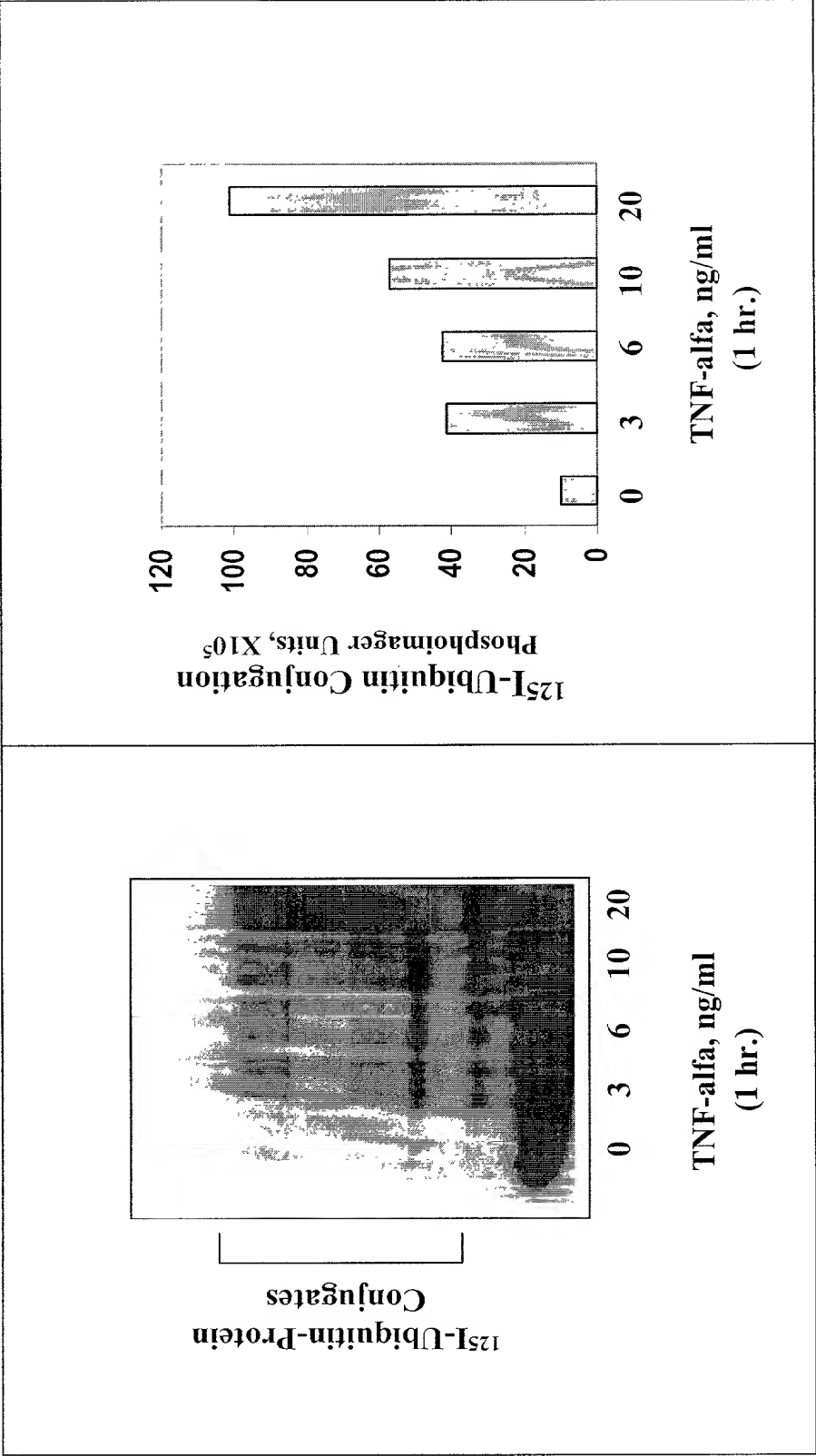


Figure 12

TNF α Elicits Accelerated Ubiquitination in C2C12 Myotube Cultures



SEQUENCE LISTING

<110> Han, Hui-Quan
Kwak, Keith

<120> Human E3 Alpha Ubiquitin Ligase

<130> 01017/35966

<140>

<141>

<150> US 60/187,911

<151> 1999-03-08

<160> 29

<170> PatentIn Ver. 2.0

<210> 1

<211> 6308

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (696)..(5942)

<400> 1

gccaaagaatt cggcacgagg ggaaaagctg agcccaggaa ccaaattact tgctttacct 60

cattgtgtaa gacaagcgtc aaaaacagct tcaacctatc ttgaacaaga gaacttacct 120

ccaaaggctt atcatctgtc ttccacttat ccaacaagct gctatggcca ctgcctgtgc 180

cgcacctgga accaccgcca gcccactac tgccctcact accactgggt ctcccacct 240

gatcagctgc ttgctgtgc catcttatcc gcttctgctt gttctgagta aatgtataca 300

caccttgga accaccattc tactttctgt gtctatgaat ttgactactc tagctggatc 360

ccgagctttt ttgtacacat gtgcaagtgc ccacggggta gaatcctaaa aatagaagat 420

gtatgcaaca gttcccagca ccaaaccag atatacaacc attcagctac caagagctac 480

gcctgataaa ttagagggga aaaaaaaat ctccagtcct ttcacgtcgt gacgcttgct 540

tccgggaagc gggccggaag ccactcctcg agtctgcgtc aaaccgcact tcaggggccc 600

tcgtaaaagt gtcgtccctg tctctccgac cggccacagg tttccgcttg cctctggccg 660

ggggtcggca actgcaggcg tcagtttccc tcaag atg gcg gac gag gag gct 713

Met Ala Asp Glu Glu Ala

1

5

gga ggt act gag agg atg gaa atc agc gcg gag tta ccc cag acc cct 761

Gly Gly Thr Glu Arg Met Glu Ile Ser Ala Glu Leu Pro Gln Thr Pro

10

15

20

cag cgt ctg gca tct tgg tgg gat cag caa gtt gat ttt tat act gct 809

Gln Arg Leu Ala Ser Trp Trp Asp Gln Gln Val Asp Phe Tyr Thr Ala

25

30

35

ttc ttg cat cat ttg gca caa ttg gtg cca gaa att tac ttt gct gaa	857
Phe Leu His His Leu Ala Gln Leu Val Pro Glu Ile Tyr Phe Ala Glu	
40 45 50	
atg gac cca gac ttg gaa aag cag gag gaa agt gta caa atg tca ata	905
Met Asp Pro Asp Leu Glu Lys Gln Glu Ser Val Gln Met Ser Ile	
55 60 65 70	
ttc act cca ctg gaa tgg tac tta ttt gga gaa gat cca gat att tgc	953
Phe Thr Pro Leu Glu Trp Tyr Leu Phe Gly Glu Asp Pro Asp Ile Cys	
75 80 85	
tta gag aaa ttg aag cac agt gga gca ttt cag ctt tgt ggg agg gtt	1001
Leu Glu Lys Leu Lys His Ser Gly Ala Phe Gln Leu Cys Gly Arg Val	
90 95 100	
ttc aaa agt gga gag aca acc tat tct tgc agg gat tgt gca att gat	1049
Phe Lys Ser Gly Glu Thr Thr Tyr Ser Cys Arg Asp Cys Ala Ile Asp	
105 110 115	
cca aca tgt gta ctc tgt atg gac tgc ttc cag gac agt gtt cat aaa	1097
Pro Thr Cys Val Leu Cys Met Asp Cys Phe Gln Asp Ser Val His Lys	
120 125 130	
aat cat cgt tac aag atg cat act tct act gga gga ggg ttc tgt gac	1145
Asn His Arg Tyr Lys Met His Thr Ser Thr Gly Gly Gly Phe Cys Asp	
135 140 145 150	
tgt gga gac aca gag gca tgg aaa act ggc cct ttt tgt gta aat cat	1193
Cys Gly Asp Thr Glu Ala Trp Lys Thr Gly Pro Phe Cys Val Asn His	
155 160 165	
gaa cct gga aga gca ggt act ata aaa gag aat tca cgc tgt ccg ttg	1241
Glu Pro Gly Arg Ala Gly Thr Ile Lys Glu Asn Ser Arg Cys Pro Leu	
170 175 180	
aat gaa gag gta att gtc caa gcc agg aaa ata ttt cct tca gtg ata	1289
Asn Glu Glu Val Ile Val Gln Ala Arg Lys Ile Phe Pro Ser Val Ile	
185 190 195	
aaa tat gtc gta gaa atg act ata tgg gaa gag gaa aaa gaa ctg cct	1337
Lys Tyr Val Val Glu Met Thr Ile Trp Glu Glu Glu Lys Glu Leu Pro	
200 205 210	
cct gaa ctc cag ata agg gag aaa aat gaa aga tac tat tgt gtc ctt	1385
Pro Glu Leu Gln Ile Arg Glu Lys Asn Glu Arg Tyr Tyr Cys Val Leu	
215 220 225 230	
ttc aat gat gaa cac cat tca tat gac cac gtc ata tac agc cta caa	1433
Phe Asn Asp Glu His His Ser Tyr Asp His Val Ile Tyr Ser Leu Gln	
235 240 245	
aga gct ctt gac tgt gag ctc gca gag gcc cag ttg cat acc act gcc	1481
Arg Ala Leu Asp Cys Glu Leu Ala Glu Ala Gln Leu His Thr Thr Ala	
250 255 260	
att gac aaa gag ggt cgt cgg gct gtt aaa gcg gga gct tat gct gct	1529
Ile Asp Lys Glu Gly Arg Arg Ala Val Lys Ala Gly Ala Tyr Ala Ala	
265 270 275	
tgc cag gaa gca aag gaa gat ata aag agt cat tca gaa aat gtc tct	1577
Cys Gln Glu Ala Lys Glu Asp Ile Lys Ser His Ser Glu Asn Val Ser	
280 285 290	

caa Gln 295	cat His	cca Pro	ctt Leu	cat His	gta Val 300	gaa Glu	gta Val	tta Leu	cac His	tca Ser 305	gag Glu	att Ile	atg Met	gct Ala	cat His 310	1625
cag Gln	aaa Lys	ttt Phe	gct Ala	ttg Leu 315	cgt Arg	ctt Leu	ggt Gly	tcc Ser	tgg Trp 320	atg Met	aac Asn	aaa Lys	att Ile	atg Met 325	agc Ser	1673
tat Tyr	tca Ser	agt Ser	gac Asp 330	ttt Phe	agg Arg	cag Gln	atc Ile	ttt Phe 335	tgc Cys	caa Gln	gca Ala	tgc Cys	ctt Leu 340	aga Arg	gaa Glu	1721
gaa Glu	cct Pro	gac Asp 345	tcg Ser	gag Glu	aat Asn	ccc Pro	tgt Cys 350	ctc Leu	ata Ile	agc Ser	agg Arg	tta Leu 355	atg Met	ctt Leu	tgg Trp	1769
gat Asp	gca Ala 360	aag Lys	ctt Leu	tat Tyr	aaa Lys	ggg Gly 365	gcc Ala	cgt Arg	aag Lys	atc Ile	ctt Leu 370	cat His	gaa Glu	ttg Leu	atc Ile	1817
ttc Phe 375	agc Ser	agt Ser	ttt Phe	ttt Phe	atg Met 380	gag Glu	atg Met	gaa Glu	tac Tyr	aaa Lys 385	aaa Lys	ctc Leu	ttt Phe	gct Ala	atg Met 390	1865
gaa Glu	ttt Phe	gtg Val	aag Lys	tat Tyr 395	tat Tyr	aaa Lys	caa Gln	ctg Leu	cag Gln 400	aaa Lys	gaa Glu	tat Tyr	atc Ile	agt Ser 405	gat Asp	1913
gat Asp	cat His	gac Asp 410	aga Arg	agt Ser	atc Ile	tct Ser	ata Ile	act Thr 415	gca Ala	ctt Leu	tca Ser	gtt Val	cag Gln 420	atg Met	ttt Phe	1961
act Thr	gtt Val	cct Pro 425	act Thr	ctg Leu	gct Ala	cga Arg	cat His 430	ctt Leu	att Ile	gaa Glu	gag Glu	cag Gln 435	aat Asn	gtt Val	atc Ile	2009
tct Ser	gtc Val 440	att Ile	act Thr	gaa Glu	act Thr	ctg Leu 445	cta Leu	gaa Glu	gtt Val	tta Leu	cct Pro 450	gag Glu	tac Tyr	ttg Leu	gac Asp	2057
agg Arg 455	aac Asn	aat Asn	aaa Lys	ttc Phe	aac Asn 460	ttc Phe	cag Gln	ggt Gly	tat Tyr	agc Ser 465	cag Gln	gac Asp	aaa Lys	ttg Leu	gga Gly 470	2105
aga Arg	gta Val	tat Tyr	gca Ala	gta Val 475	ata Ile	tgt Cys	gac Asp	cta Leu	aag Lys 480	tat Tyr	atc Ile	ctg Leu	atc Ile	agc Ser 485	aaa Lys	2153
ccc Pro	aca Thr	ata Ile	tgg Trp 490	aca Thr	gaa Glu	aga Arg	tta Leu	aga Arg 495	atg Met	cag Gln	ttc Phe	ctt Leu	gaa Glu 500	ggg Gly	ttt Phe	2201
cga Arg	tct Ser	ttt Phe 505	ttg Leu	aag Lys	att Ile	ctt Leu	acc Thr 510	tgt Cys	atg Met	cag Gln	gga Gly	atg Met 515	gaa Glu	gaa Glu	atc Ile	2249
cga Arg	aga Arg 520	cag Gln	gtt Val	ggg Gly	caa Gln	cac His 525	att Ile	gaa Glu	gtg Val	gat Asp 530	cct Pro	gat Asp	tgg Trp	gag Glu	gct Ala	2297
gcc Ala 535	att Ile	gct Ala	ata Ile	cag Gln	atg Met 540	caa Gln	ttg Leu	aag Lys	aat Asn 545	att Ile	tta Leu	ctc Leu	atg Met	ttc Phe	caa Gln 550	2345

gag tgg tgt gct tgt gat gaa gaa ctc tta ctt gtg gct tat aaa gaa	2393
Glu Trp Cys Ala Cys Asp Glu Glu Leu Leu Leu Val Ala Tyr Lys Glu	
555 560 565	
tgt cac aaa gct gtg atg agg tgc agt acc agt ttc ata tct agt agc	2441
Cys His Lys Ala Val Met Arg Cys Ser Thr Ser Phe Ile Ser Ser Ser	
570 575 580	
aag aca gta gta caa tcg tgt gga cat agt ttg gaa aca aag tcc tac	2489
Lys Thr Val Val Gln Ser Cys Gly His Ser Leu Glu Thr Lys Ser Tyr	
585 590 595	
aga gta tct gag gat ctt gta agc ata cat ctg cca ctc tct agg acc	2537
Arg Val Ser Glu Asp Leu Val Ser Ile His Leu Pro Leu Ser Arg Thr	
600 605 610	
ctt gct ggt ctt cat gta cgt tta agc agg ctg ggt gct gtt tca aga	2585
Leu Ala Gly Leu His Val Arg Leu Ser Arg Leu Gly Ala Val Ser Arg	
615 620 625 630	
ctg cat gaa ttt gtg tct ttt gag gac ttt caa gta gag gta cta gtg	2633
Leu His Glu Phe Val Ser Phe Glu Asp Phe Gln Val Glu Val Leu Val	
635 640 645	
gaa tat cct tta cgt tgt ctg gtg ttg gtt gcc cag gtt gtt gct gag	2681
Glu Tyr Pro Leu Arg Cys Leu Val Leu Val Ala Gln Val Val Ala Glu	
650 655 660	
atg tgg cga aga aat gga ctg tct ctt att agc cag gtg ttt tat tac	2729
Met Trp Arg Arg Asn Gly Leu Ser Leu Ile Ser Gln Val Phe Tyr Tyr	
665 670 675	
caa gat gtt aag tgc aga gaa gaa atg tat gat aaa gat atc atc atg	2777
Gln Asp Val Lys Cys Arg Glu Glu Met Tyr Asp Lys Asp Ile Ile Met	
680 685 690	
ctt cag att ggt gca tct tta atg gat ccc aat aag ttc ttg tta ctg	2825
Leu Gln Ile Gly Ala Ser Leu Met Asp Pro Asn Lys Phe Leu Leu Leu	
695 700 705 710	
gta ctt cag agg tat gaa ctt gcc gag gct ttt aac aag acc ata tct	2873
Val Leu Gln Arg Tyr Glu Leu Ala Glu Ala Phe Asn Lys Thr Ile Ser	
715 720 725	
aca aaa gac cag gat ttg att aaa caa tat aat aca cta ata gaa gaa	2921
Thr Lys Asp Gln Asp Leu Ile Lys Gln Tyr Asn Thr Leu Ile Glu Glu	
730 735 740	
atg ctt cag gtc ctc atc tat att gtg ggt gag cgt tat gta cct gga	2969
Met Leu Gln Val Leu Ile Tyr Ile Val Gly Glu Arg Tyr Val Pro Gly	
745 750 755	
gtg gga aat gtg acc aaa gaa gag gtc aca atg aga gaa atc att cac	3017
Val Gly Asn Val Thr Lys Glu Glu Val Thr Met Arg Glu Ile Ile His	
760 765 770	
ttg ctt tgc att gaa ccc atg cca cac agt gcc att gcc aaa aat tta	3065
Leu Leu Cys Ile Glu Pro Met Pro His Ser Ala Ile Ala Lys Asn Leu	
775 780 785 790	
cct gag aat gaa aat aat gaa act ggc tta gag aat gtc ata aac aaa	3113
Pro Glu Asn Glu Asn Asn Glu Thr Gly Leu Glu Asn Val Ile Asn Lys	
795 800 805	

gtg	gcc	aca	ttt	aag	aaa	cca	ggg	gta	tca	ggc	cat	gga	gtt	tat	gaa	3161
Val	Ala	Thr	Phe	Lys	Lys	Pro	Gly	Val	Ser	Gly	His	Gly	Val	Tyr	Glu	
			810				815						820			
cta	aaa	gat	gaa	tca	ctg	aaa	gac	ttc	aat	atg	tac	ttt	tat	cat	tac	3209
Leu	Lys	Asp	Glu	Ser	Leu	Lys	Asp	Phe	Asn	Met	Tyr	Phe	Tyr	His	Tyr	
			825				830						835			
tcc	aaa	acc	cag	cat	agc	aag	gct	gaa	cat	atg	cag	aag	aaa	agg	aga	3257
Ser	Lys	Thr	Gln	His	Ser	Lys	Ala	Glu	His	Met	Gln	Lys	Lys	Arg	Arg	
			840				845						850			
aaa	caa	gaa	aac	aaa	gat	gaa	gca	ttg	ccg	cca	cca	cca	cct	cct	gaa	3305
Lys	Gln	Glu	Asn	Lys	Asp	Glu	Ala	Leu	Pro	Pro	Pro	Pro	Pro	Pro	Glu	
			855				860						865			870
ttc	tgc	cct	gct	ttc	agc	aaa	gtg	att	aac	ctt	ctc	aac	tgt	gat	atc	3353
Phe	Cys	Pro	Ala	Phe	Ser	Lys	Val	Ile	Asn	Leu	Leu	Asn	Cys	Asp	Ile	
			875				880						885			
atg	atg	tac	att	ctc	agg	acc	gta	ttt	gag	cgg	gca	ata	gac	aca	gat	3401
Met	Met	Tyr	Ile	Leu	Arg	Thr	Val	Phe	Glu	Arg	Ala	Ile	Asp	Thr	Asp	
			890				895						900			
tct	aac	ttg	tgg	acc	gaa	ggg	atg	ctc	caa	atg	gct	ttt	cat	att	ctg	3449
Ser	Asn	Leu	Trp	Thr	Glu	Gly	Met	Leu	Gln	Met	Ala	Phe	His	Ile	Leu	
			905				910						915			
gca	ttg	ggg	tta	cta	gaa	gag	aag	caa	cag	ctt	caa	aaa	gct	cct	gaa	3497
Ala	Leu	Gly	Leu	Leu	Glu	Glu	Lys	Gln	Gln	Leu	Gln	Lys	Ala	Pro	Glu	
			920				925						930			
gaa	gaa	gta	aca	ttt	gac	ttt	tat	cat	aag	gct	tca	aga	ttg	gga	agt	3545
Glu	Glu	Val	Thr	Phe	Asp	Phe	Tyr	His	Lys	Ala	Ser	Arg	Leu	Gly	Ser	
			935				940						945			950
tca	gcc	atg	aat	ata	caa	atg	ctt	ttg	gaa	aaa	ctc	aaa	gga	att	ccc	3593
Ser	Ala	Met	Asn	Ile	Gln	Met	Leu	Leu	Glu	Lys	Leu	Lys	Gly	Ile	Pro	
			955				960						965			
cag	tta	gaa	ggc	cag	aag	gac	atg	ata	acg	tgg	ata	ctt	cag	atg	ttt	3641
Gln	Leu	Glu	Gly	Gln	Lys	Asp	Met	Ile	Thr	Trp	Ile	Leu	Gln	Met	Phe	
			970				975						980			
gac	aca	gtg	aag	cga	tta	aga	gaa	aaa	tct	tgt	tta	att	gta	gca	acc	3689
Asp	Thr	Val	Lys	Arg	Leu	Arg	Glu	Lys	Ser	Cys	Leu	Ile	Val	Ala	Thr	
			985				990						995			
aca	tca	gga	tcg	gaa	tct	att	aag	aat	gat	gag	att	act	cat	gat	aaa	3737
Thr	Ser	Gly	Ser	Glu	Ser	Ile	Lys	Asn	Asp	Glu	Ile	Thr	His	Asp	Lys	
			1000				1005						1010			
gaa	aaa	gca	gaa	cga	aaa	aga	aaa	gct	gaa	gct	gct	agg	cta	cat	cgc	3785
Glu	Lys	Ala	Glu	Arg	Lys	Arg	Lys	Ala	Glu	Ala	Ala	Arg	Leu	His	Arg	
			1015				1020						1025			1030
cag	aag	atc	atg	gct	cag	atg	tct	gcc	tta	cag	aaa	aac	ttc	att	gaa	3833
Gln																

gat tcc att atg gag gaa gag agc acc cca gca gtc agt gac tac tct	3929
Asp Ser Ile Met Glu Glu Glu Ser Thr Pro Ala Val Ser Asp Tyr Ser	
1065 1070 1075	
aga att gct ttg ggt cct aaa cgg ggt cca tct gtt act gaa aag gag	3977
Arg Ile Ala Leu Gly Pro Lys Arg Gly Pro Ser Val Thr Glu Lys Glu	
1080 1085 1090	
gtg ctg acg tgc atc ctt tgc caa gaa gaa cag gag gtg aaa ata gaa	4025
Val Leu Thr Cys Ile Leu Cys Gln Glu Glu Gln Glu Val Lys Ile Glu	
1095 1100 1105 1110	
aat aat gcc atg gta tta tgc gcc tgt gtc cag aaa tct act gcc tta	4073
Asn Asn Ala Met Val Leu Ser Ala Cys Val Gln Lys Ser Thr Ala Leu	
1115 1120 1125	
acc cag cac agg gga aaa ccc ata gaa ctc tca gga gaa gcc cta gac	4121
Thr Gln His Arg Gly Lys Pro Ile Glu Leu Ser Gly Glu Ala Leu Asp	
1130 1135 1140	
cca ctt ttc atg gat cca gac ttg gca tat gga act tat aca gga agc	4169
Pro Leu Phe Met Asp Pro Asp Leu Ala Tyr Gly Thr Tyr Thr Gly Ser	
1145 1150 1155	
tgt ggt cat gta atg cac gca gtg tgc tgg cag aag tat ttt gaa gct	4217
Cys Gly His Val Met His Ala Val Cys Trp Gln Lys Tyr Phe Glu Ala	
1160 1165 1170	
gta cag ctg agc tct cag cag cgc att cat gtt gac ctt ttt gac ttg	4265
Val Gln Leu Ser Ser Gln Gln Arg Ile His Val Asp Leu Phe Asp Leu	
1175 1180 1185 1190	
gaa agt gga gaa tat ctt tgc cct ctt tgc aaa tct ctg tgc aat act	4313
Glu Ser Gly Glu Tyr Leu Cys Pro Leu Cys Lys Ser Leu Cys Asn Thr	
1195 1200 1205	
gtg atc ccc att att cct ttg caa cct caa aag ata aac agt gag aat	4361
Val Ile Pro Ile Ile Pro Leu Gln Pro Gln Lys Ile Asn Ser Glu Asn	
1210 1215 1220	
gca gat gct ctt gct caa ctt ttg acc ctg gca cgg tgg ata cag act	4409
Ala Asp Ala Leu Ala Gln Leu Leu Thr Leu Ala Arg Trp Ile Gln Thr	
1225 1230 1235	
gtt ctg gcc aga ata tca ggt tat aat ata aga cat gct aaa gga gaa	4457
Val Leu Ala Arg Ile Ser Gly Tyr Asn Ile Arg His Ala Lys Gly Glu	
1240 1245 1250	
aac cca att cct att ttc ttt aat caa gga atg gga gat tct act ttg	4505
Asn Pro Ile Pro Ile Phe Phe Asn Gln Gly Met Gly Asp Ser Thr Leu	
1255 1260 1265 1270	
gag ttc cat tcc atc ctg agt ttt ggc gtt gag tct tgc att aaa tat	4553
Glu Phe His Ser Ile Leu Ser Phe Gly Val Glu Ser Ser Ile Lys Tyr	
1275 1280 1285	
tca aat agc atc aag gaa atg gtt att ctc ttt gcc aca aca att tat	4601
Ser Asn Ser Ile Lys Glu Met Val Ile Leu Phe Ala Thr Thr Ile Tyr	
1290 1295 1300	
aga att gga ttg aaa gtg cca cct gat gaa agg gat cct cga gtc ccc	4649
Arg Ile Gly Leu Lys Val Pro Pro Asp Glu Arg Asp Pro Arg Val Pro	
1305 1310 1315	

atg ctg acc tgg agc acc tgc gct ttc act atc cag gca att gaa aat	4697
Met Leu Thr Trp Ser Thr Cys Ala Phe Thr Ile Gln Ala Ile Glu Asn	
1320 1325 1330	
cta ttg gga gat gaa gga aaa cct ctg ttt gga gca ctt caa aat agg	4745
Leu Leu Gly Asp Glu Gly Lys Pro Leu Phe Gly Ala Leu Gln Asn Arg	
1335 1340 1345 1350	
cag cat aat ggt ctg aaa gca tta atg cag ttt gca gtt gca cag agg	4793
Gln His Asn Gly Leu Lys Ala Leu Met Gln Phe Ala Val Ala Gln Arg	
1355 1360 1365	
att acc tgt cct cag gtc ctg ata cag aaa cat ctg gtt cgt ctt cta	4841
Ile Thr Cys Pro Gln Val Leu Ile Gln Lys His Leu Val Arg Leu Leu	
1370 1375 1380	
tca gtt gtt ctt cct aac ata aaa tca gaa gat aca cca tgc ctt ctg	4889
Ser Val Val Leu Pro Asn Ile Lys Ser Glu Asp Thr Pro Cys Leu Leu	
1385 1390 1395	
tct ata gat ctg ttt cat gtt ttg gtg ggt gct gtg tta gca ttc cca	4937
Ser Ile Asp Leu Phe His Val Leu Val Gly Ala Val Leu Ala Phe Pro	
1400 1405 1410	
tcc ttg tat tgg gat gac cct gtt gat ctg cag cct tct tca gtt agt	4985
Ser Leu Tyr Trp Asp Asp Pro Val Asp Leu Gln Pro Ser Ser Val Ser	
1415 1420 1425 1430	
tct tcc tat aac cac ctt tat ctc ttc cat ttg atc acc atg gca cac	5033
Ser Ser Tyr Asn His Leu Tyr Leu Phe His Leu Ile Thr Met Ala His	
1435 1440 1445	
atg ctt cag ata cta ctt aca gta gac aca ggc cta ccc ctt gct cag	5081
Met Leu Gln Ile Leu Leu Thr Val Asp Thr Gly Leu Pro Leu Ala Gln	
1450 1455 1460	
gtt caa gaa gac agt gaa gag gct cat tcc gca tct tct ttc ttt gca	5129
Val Gln Glu Asp Ser Glu Glu Ala His Ser Ala Ser Ser Phe Phe Ala	
1465 1470 1475	
gaa att tct caa tat aca agt ggc tcc att ggg tgt gat att cct ggc	5177
Glu Ile Ser Gln Tyr Thr Ser Gly Ser Ile Gly Cys Asp Ile Pro Gly	
1480 1485 1490	
tgg tat ttg tgg gtc tca ctg aag aat ggc atc acc cct tat ctt cgc	5225
Trp Tyr Leu Trp Val Ser Leu Lys Asn Gly Ile Thr Pro Tyr Leu Arg	
1495 1500 1505 1510	
tgt gct gca ttg ttt ttc cac tat tta ctt ggg gta act ccg cct gag	5273
Cys Ala Ala Leu Phe Phe His Tyr Leu Leu Gly Val Thr Pro Pro Glu	
1515 1520 1525	
gaa ctg cat acc aat tct gca gaa gga gag tac agt gca ctc tgt agc	5321
Glu Leu His Thr Asn Ser Ala Glu Gly Glu Tyr Ser Ala Leu Cys Ser	
1530 1535 1540	
tat cta tct tta cct aca aat ttg ttc ctg ctc ttc cag gaa tat tgg	5369
Tyr Leu Ser Leu Pro Thr Asn Leu Phe Leu Leu Phe Gln Glu Tyr Trp	
1545 1550 1555	
gat act gta agg ccc ttg ctc cag agg tgg tgt gca gat cct gcc tta	5417
Asp Thr Val Arg Pro Leu Leu Gln Arg Trp Cys Ala Asp Pro Ala Leu	
1560 1565 1570	

cta aac tgt ttg aag caa aaa aac acc gtg gtc agg tac cct aga aaa 5465
 Leu Asn Cys Leu Lys Gln Lys Asn Thr Val Val Arg Tyr Pro Arg Lys
 1575 1580 1585 1590

aga aat agt ttg ata gag ctt cct gat gac tat agc tgc ctc ctg aat 5513
 Arg Asn Ser Leu Ile Glu Leu Pro Asp Asp Tyr Ser Cys Leu Leu Asn
 1595 1600 1605

caa gct tct cat ttc agg tgc cca cgg tct gca gat gat gag cga aag 5561
 Gln Ala Ser His Phe Arg Cys Pro Arg Ser Ala Asp Asp Glu Arg Lys
 1610 1615 1620

cat cct gtc ctc tgc ctt ttc tgt ggg gct ata cta tgt tct cag aac 5609
 His Pro Val Leu Cys Leu Phe Cys Gly Ala Ile Leu Cys Ser Gln Asn
 1625 1630 1635

att tgc tgc cag gaa att gtg aac ggg gaa gag gtt gga gct tgc att 5657
 Ile Cys Cys Gln Glu Ile Val Asn Gly Glu Glu Val Gly Ala Cys Ile
 1640 1645 1650

ttt cac gca ctt cac tgt gga gcc gga gtc tgc att ttc cta aaa atc 5705
 Phe His Ala Leu His Cys Gly Ala Gly Val Cys Ile Phe Leu Lys Ile
 1655 1660 1665 1670

aga gaa tgc cga gtg gtc ctg gtt gaa ggt aaa gcc aga ggc tgt gcc 5753
 Arg Glu Cys Arg Val Val Leu Val Glu Gly Lys Ala Arg Gly Cys Ala
 1675 1680 1685

tat cca gct cct tac ttg gat gaa tat gga gaa aca gac cct ggc ctg 5801
 Tyr Pro Ala Pro Tyr Leu Asp Glu Tyr Gly Glu Thr Asp Pro Gly Leu
 1690 1695 1700

aag agg ggc aac ccc ctt cat tta tct cgt gag cgg tat cgg aag ctc 5849
 Lys Arg Gly Asn Pro Leu His Leu Ser Arg Glu Arg Tyr Arg Lys Leu
 1705 1710 1715

cat ttg gtc tgg caa caa cac tgc att ata gaa gag att gct agg agc 5897
 His Leu Val Trp Gln Gln His Cys Ile Ile Glu Glu Ile Ala Arg Ser
 1720 1725 1730

caa gag act aat cag atg tta ttt gga ttc aac tgg cag tta ctg 5942
 Gln Glu Thr Asn Gln Met Leu Phe Gly Phe Asn Trp Gln Leu Leu
 1735 1740 1745

tgagctccaa ctctgcctca agacaatcac aaatgacgac agtagtaaag gctgattcaa 6002

aattatggaa aacttttctga gggctgggaa agtattggag ggtcttttgc tccatgtcca 6062

ggttcactta catcaataaaa atattttctta atggagtatt gcttttcaatt agcaaacata 6122

tgcttcacag gaaaaaagga catagatcaa tctgttttat gtgctagtat ttccaggaat 6182

ttattcccct tcataatttg tctcatttca ttttatttca tccacttggt agatgaagtc 6242

acgtcaaaca gttgtagaca ttttatgtgt tggttaactc ttctgcaatt ttgtatttgg 6302

tggtttt 6308

<210> 2
 <211> 1749
 <212> PRT
 <213> Homo sapiens

Met 1	Ala	Asp	Glu	Glu 5	Ala	Gly	Gly	Thr	Glu 10	Arg	Met	Glu	Ile	Ser 15	Ala
Glu	Leu	Pro	Gln 20	Thr	Pro	Gln	Arg	Leu 25	Ala	Ser	Trp	Trp	Asp 30	Gln	Gln
Val	Asp	Phe 35	Tyr	Thr	Ala	Phe	Leu 40	His	His	Leu	Ala	Gln 45	Leu	Val	Pro
Glu	Ile 50	Tyr	Phe	Ala	Glu	Met 55	Asp	Pro	Asp	Leu	Glu 60	Lys	Gln	Glu	Glu
Ser 65	Val	Gln	Met	Ser	Ile 70	Phe	Thr	Pro	Leu	Glu 75	Trp	Tyr	Leu	Phe	Gly 80
Glu	Asp	Pro	Asp	Ile 85	Cys	Leu	Glu	Lys	Leu 90	Lys	His	Ser	Gly	Ala 95	Phe
Gln	Leu	Cys	Gly 100	Arg	Val	Phe	Lys	Ser 105	Gly	Glu	Thr	Thr	Tyr	Ser	Cys
Arg	Asp	Cys 115	Ala	Ile	Asp	Pro	Thr 120	Cys	Val	Leu	Cys	Met 125	Asp	Cys	Phe
Gln	Asp 130	Ser	Val	His	Lys	Asn 135	His	Arg	Tyr	Lys	Met 140	His	Thr	Ser	Thr
Gly 145	Gly	Gly	Phe	Cys 150	Asp	Cys	Gly	Asp	Thr	Glu 155	Ala	Trp	Lys	Thr	Gly 160
Pro	Phe	Cys	Val	Asn 165	His	Glu	Pro	Gly	Arg 170	Ala	Gly	Thr	Ile	Lys 175	Glu
Asn	Ser	Arg	Cys 180	Pro	Leu	Asn	Glu	Glu 185	Val	Ile	Val	Gln	Ala 190	Arg	Lys
Ile	Phe	Pro 195	Ser	Val	Ile	Lys	Tyr 200	Val	Val	Glu	Met 205	Thr	Ile	Trp	Glu
Glu 210	Glu	Lys	Glu	Leu	Pro	Pro 215	Glu	Leu	Gln	Ile	Arg 220	Glu	Lys	Asn	Glu
Arg 225	Tyr	Tyr	Cys	Val	Leu 230	Phe	Asn	Asp	Glu	His 235	His	Ser	Tyr	Asp	His 240
Val	Ile	Tyr	Ser	Leu 245	Gln	Arg	Ala	Leu	Asp 250	Cys	Glu	Leu	Ala	Glu 255	Ala
Gln	Leu	His	Thr 260	Thr	Ala	Ile	Asp	Lys 265	Glu	Gly	Arg	Arg	Ala 270	Val	Lys
Ala	Gly	Ala 275	Tyr	Ala	Ala	Cys	Gln 280	Glu	Ala	Lys	Glu	Asp 285	Ile	Lys	Ser
His 290	Ser	Glu	Asn	Val	Ser	Gln 295	His	Pro	Leu	His	Val 300	Glu	Val	Leu	His
Ser 305	Glu	Ile	Met	Ala	His 310	Gln	Lys	Phe	Ala	Leu 315	Arg	Leu	Gly	Ser	Trp 320
Met	Asn	Lys	Ile	Met 325	Ser	Tyr	Ser	Ser	Asp 330	Phe	Arg	Gln	Ile	Phe 335	Cys

Gln	Ala	Cys	Leu	Arg	Glu	Glu	Pro	Asp	Ser	Glu	Asn	Pro	Cys	Leu	Ile	340	345	350
Ser	Arg	Leu	Met	Leu	Trp	Asp	Ala	Lys	Leu	Tyr	Lys	Gly	Ala	Arg	Lys	355	360	365
Ile	Leu	His	Glu	Leu	Ile	Phe	Ser	Ser	Phe	Phe	Met	Glu	Met	Glu	Tyr	370	375	380
Lys	Lys	Leu	Phe	Ala	Met	Glu	Phe	Val	Lys	Tyr	Tyr	Lys	Gln	Leu	Gln	385	390	400
Lys	Glu	Tyr	Ile	Ser	Asp	Asp	His	Asp	Arg	Ser	Ile	Ser	Ile	Thr	Ala	405	410	415
Leu	Ser	Val	Gln	Met	Phe	Thr	Val	Pro	Thr	Leu	Ala	Arg	His	Leu	Ile	420	425	430
Glu	Glu	Gln	Asn	Val	Ile	Ser	Val	Ile	Thr	Glu	Thr	Leu	Leu	Glu	Val	435	440	445
Leu	Pro	Glu	Tyr	Leu	Asp	Arg	Asn	Asn	Lys	Phe	Asn	Phe	Gln	Gly	Tyr	450	455	460
Ser	Gln	Asp	Lys	Leu	Gly	Arg	Val	Tyr	Ala	Val	Ile	Cys	Asp	Leu	Lys	465	470	475
Tyr	Ile	Leu	Ile	Ser	Lys	Pro	Thr	Ile	Trp	Thr	Glu	Arg	Leu	Arg	Met	485	490	495
Gln	Phe	Leu	Glu	Gly	Phe	Arg	Ser	Phe	Leu	Lys	Ile	Leu	Thr	Cys	Met	500	505	510
Gln	Gly	Met	Glu	Glu	Ile	Arg	Arg	Gln	Val	Gly	Gln	His	Ile	Glu	Val	515	520	525
Asp	Pro	Asp	Trp	Glu	Ala	Ala	Ile	Ala	Ile	Gln	Met	Gln	Leu	Lys	Asn	530	535	540
Ile	Leu	Leu	Met	Phe	Gln	Glu	Trp	Cys	Ala	Cys	Asp	Glu	Glu	Leu	Leu	545	550	555
Leu	Val	Ala	Tyr	Lys	Glu	Cys	His	Lys	Ala	Val	Met	Arg	Cys	Ser	Thr	565	570	575
Ser	Phe	Ile	Ser	Ser	Ser	Lys	Thr	Val	Val	Gln	Ser	Cys	Gly	His	Ser	580	585	590
Leu	Glu	Thr	Lys	Ser	Tyr	Arg	Val	Ser	Glu	Asp	Leu	Val	Ser	Ile	His	595	600	605
Leu	Pro	Leu	Ser	Arg	Thr	Leu	Ala	Gly	Leu	His	Val	Arg	Leu	Ser	Arg	610	615	620
Leu	Gly	Ala	Val	Ser	Arg	Leu	His	Glu	Phe	Val	Ser	Phe	Glu	Asp	Phe	625	630	635
Gln	Val	Glu	Val	Leu	Val	Glu	Tyr	Pro	Leu	Arg	Cys	Leu	Val	Leu	Val	645	650	655
Ala	Gln	Val	Val	Ala	Glu	Met	Trp	Arg	Arg	Asn	Gly	Leu	Ser	Leu	Ile	660	665	670

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

Ser	Gln	Val	Phe	Tyr	Tyr	Gln	Asp	Val	Lys	Cys	Arg	Glu	Glu	Met	Tyr
675						680						685			
Asp	Lys	Asp	Ile	Ile	Met	Leu	Gln	Ile	Gly	Ala	Ser	Leu	Met	Asp	Pro
690						695				700					
Asn	Lys	Phe	Leu	Leu	Leu	Val	Leu	Gln	Arg	Tyr	Glu	Leu	Ala	Glu	Ala
705				710						715				720	
Phe	Asn	Lys	Thr	Ile	Ser	Thr	Lys	Asp	Gln	Asp	Leu	Ile	Lys	Gln	Tyr
				725				730						735	
Asn	Thr	Leu	Ile	Glu	Glu	Met	Leu	Gln	Val	Leu	Ile	Tyr	Ile	Val	Gly
		740						745				750			
Glu	Arg	Tyr	Val	Pro	Gly	Val	Gly	Asn	Val	Thr	Lys	Glu	Glu	Val	Thr
		755				760						765			
Met	Arg	Glu	Ile	Ile	His	Leu	Leu	Cys	Ile	Glu	Pro	Met	Pro	His	Ser
770						775				780					
Ala	Ile	Ala	Lys	Asn	Leu	Pro	Glu	Asn	Glu	Asn	Asn	Glu	Thr	Gly	Leu
785				790						795				800	
Glu	Asn	Val	Ile	Asn	Lys	Val	Ala	Thr	Phe	Lys	Lys	Pro	Gly	Val	Ser
				805				810						815	
Gly	His	Gly	Val	Tyr	Glu	Leu	Lys	Asp	Glu	Ser	Leu	Lys	Asp	Phe	Asn
		820						825				830			
Met	Tyr	Phe	Tyr	His	Tyr	Ser	Lys	Thr	Gln	His	Ser	Lys	Ala	Glu	His
835						840				845					
Met	Gln	Lys	Lys	Arg	Arg	Lys	Gln	Glu	Asn	Lys	Asp	Glu	Ala	Leu	Pro
850						855				860					
Pro	Pro	Pro	Pro	Pro	Glu	Phe	Cys	Pro	Ala	Phe	Ser	Lys	Val	Ile	Asn
865				870						875				880	
Leu	Leu	Asn	Cys	Asp	Ile	Met	Met	Tyr	Ile	Leu	Arg	Thr	Val	Phe	Glu
				885				890						895	
Arg	Ala	Ile	Asp	Thr	Asp	Ser	Asn	Leu	Trp	Thr	Glu	Gly	Met	Leu	Gln
		900						905				910			
Met	Ala	Phe	His	Ile	Leu	Ala	Leu	Gly	Leu	Leu	Glu	Glu	Lys	Gln	Gln
915						920						925			
Leu	Gln	Lys	Ala	Pro	Glu	Glu	Glu	Val	Thr	Phe	Asp	Phe	Tyr	His	Lys
930						935				940					
Ala	Ser	Arg	Leu	Gly	Ser	Ser	Ala	Met	Asn	Ile	Gln	Met	Leu	Leu	Glu
945				950						955				960	
Lys	Leu	Lys	Gly	Ile	Pro	Gln	Leu	Glu	Gly	Gln	Lys	Asp	Met	Ile	Thr
				965				970						975	
Trp	Ile	Leu	Gln	Met	Phe	Asp	Thr	Val	Lys	Arg	Leu	Arg	Glu	Lys	Ser
		980						985				990			
Cys	Leu	Ile	Val	Ala	Thr	Thr	Ser	Gly	Ser	Glu	Ser	Ile	Lys	Asn	Asp
995						1000						1005			

Glu Ile Thr His Asp Lys Glu Lys Ala Glu Arg Lys Arg Lys Ala Glu
 1010 1015 1020
 Ala Ala Arg Leu His Arg Gln Lys Ile Met Ala Gln Met Ser Ala Leu
 025 1030 1035 1040
 Gln Lys Asn Phe Ile Glu Thr His Lys Leu Met Tyr Asp Asn Thr Ser
 1045 1050 1055
 Glu Met Pro Gly Lys Glu Asp Ser Ile Met Glu Glu Glu Ser Thr Pro
 1060 1065 1070
 Ala Val Ser Asp Tyr Ser Arg Ile Ala Leu Gly Pro Lys Arg Gly Pro
 1075 1080 1085
 Ser Val Thr Glu Lys Glu Val Leu Thr Cys Ile Leu Cys Gln Glu Glu
 1090 1095 1100
 Gln Glu Val Lys Ile Glu Asn Asn Ala Met Val Leu Ser Ala Cys Val
 1105 1110 1115 1120
 Gln Lys Ser Thr Ala Leu Thr Gln His Arg Gly Lys Pro Ile Glu Leu
 1125 1130 1135
 Ser Gly Glu Ala Leu Asp Pro Leu Phe Met Asp Pro Asp Leu Ala Tyr
 1140 1145 1150
 Gly Thr Tyr Thr Gly Ser Cys Gly His Val Met His Ala Val Cys Trp
 1155 1160 1165
 Gln Lys Tyr Phe Glu Ala Val Gln Leu Ser Ser Gln Gln Arg Ile His
 1170 1175 1180
 Val Asp Leu Phe Asp Leu Glu Ser Gly Glu Tyr Leu Cys Pro Leu Cys
 1185 1190 1195 1200
 Lys Ser Leu Cys Asn Thr Val Ile Pro Ile Ile Pro Leu Gln Pro Gln
 1205 1210 1215
 Lys Ile Asn Ser Glu Asn Ala Asp Ala Leu Ala Gln Leu Leu Thr Leu
 1220 1225 1230
 Ala Arg Trp Ile Gln Thr Val Leu Ala Arg Ile Ser Gly Tyr Asn Ile
 1235 1240 1245
 Arg His Ala Lys Gly Glu Asn Pro Ile Pro Ile Phe Phe Asn Gln Gly
 1250 1255 1260
 Met Gly Asp Ser Thr Leu Glu Phe His Ser Ile Leu Ser Phe Gly Val
 1265 1270 1275 1280
 Glu Ser Ser Ile Lys Tyr Ser Asn Ser Ile Lys Glu Met Val Ile Leu
 1285 1290 1295
 Phe Ala Thr Thr Ile Tyr Arg Ile Gly Leu Lys Val Pro Pro Asp Glu
 1300 1305 1310
 Arg Asp Pro Arg Val Pro Met Leu Thr Trp Ser Thr Cys Ala Phe Thr
 1315 1320 1325
 Ile Gln Ala Ile Glu Asn Leu Leu Gly Asp Glu Gly Lys Pro Leu Phe
 1330 1335 1340

Gly Ala Leu Gln Asn Arg Gln His Asn Gly Leu Lys Ala Leu Met Gln
345 1350 1355 1360

Phe Ala Val Ala Gln Arg Ile Thr Cys Pro Gln Val Leu Ile Gln Lys
1365 1370 1375

His Leu Val Arg Leu Leu Ser Val Val Leu Pro Asn Ile Lys Ser Glu
1380 1385 1390

Asp Thr Pro Cys Leu Leu Ser Ile Asp Leu Phe His Val Leu Val Gly
1395 1400 1405

Ala Val Leu Ala Phe Pro Ser Leu Tyr Trp Asp Asp Pro Val Asp Leu
1410 1415 1420

Gln Pro Ser Ser Val Ser Ser Tyr Asn His Leu Tyr Leu Phe His
425 1430 1435 1440

Leu Ile Thr Met Ala His Met Leu Gln Ile Leu Leu Thr Val Asp Thr
1445 1450 1455

Gly Leu Pro Leu Ala Gln Val Gln Glu Asp Ser Glu Glu Ala His Ser
1460 1465 1470

Ala Ser Ser Phe Phe Ala Glu Ile Ser Gln Tyr Thr Ser Gly Ser Ile
1475 1480 1485

Gly Cys Asp Ile Pro Gly Trp Tyr Leu Trp Val Ser Leu Lys Asn Gly
1490 1495 1500

Ile Thr Pro Tyr Leu Arg Cys Ala Ala Leu Phe Phe His Tyr Leu Leu
505 1510 1515 1520

Gly Val Thr Pro Pro Glu Glu Leu His Thr Asn Ser Ala Glu Gly Glu
1525 1530 1535

Tyr Ser Ala Leu Cys Ser Tyr Leu Ser Leu Pro Thr Asn Leu Phe Leu
1540 1545 1550

Leu Phe Gln Glu Tyr Trp Asp Thr Val Arg Pro Leu Leu Gln Arg Trp
1555 1560 1565

Cys Ala Asp Pro Ala Leu Leu Asn Cys Leu Lys Gln Lys Asn Thr Val
1570 1575 1580

Val Arg Tyr Pro Arg Lys Arg Asn Ser Leu Ile Glu Leu Pro Asp Asp
585 1590 1595 1600

Tyr Ser Cys Leu Leu Asn Gln Ala Ser His Phe Arg Cys Pro Arg Ser
1605 1610 1615

Ala Asp Asp Glu Arg Lys His Pro Val Leu Cys Leu Phe Cys Gly Ala
1620 1625 1630

Ile Leu Cys Ser Gln Asn Ile Cys Cys Gln Glu Ile Val Asn Gly Glu
1635 1640 1645

Glu Val Gly Ala Cys Ile Phe His Ala Leu His Cys Gly Ala Gly Val
1650 1655 1660

Cys Ile Phe Leu Lys Ile Arg Glu Cys Arg Val Val Leu Val Glu Gly
665 1670 1675 1680

1000 900 800 700 600 500 400 300 200 100 0

Lys Ala Arg Gly Cys Ala Tyr Pro Ala Pro Tyr Leu Asp Glu Tyr Gly
1685 1690 1695

Glu Thr Asp Pro Gly Leu Lys Arg Gly Asn Pro Leu His Leu Ser Arg
1700 1705 1710

Glu Arg Tyr Arg Lys Leu His Leu Val Trp Gln Gln His Cys Ile Ile
1715 1720 1725

Glu Glu Ile Ala Arg Ser Gln Glu Thr Asn Gln Met Leu Phe Gly Phe
1730 1735 1740

Asn Trp Gln Leu Leu
745

<210> 3
<211> 6300
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (295) .. (5559)

<400> 3
gccaaagaatt cggcacgagg tgtcaggcct ggggttttct gtgtccttcc ctgggtcagg 60
gacgagccag tgacttgact cttgggcgct aagcttggga gggagcgcag gaggccgctg 120
tccttccttt ccggttcacg tcacctttct ctccctctgt tgctccacct gcagccactt 180
ggacggctcc gggactgatt gcctggggca ggggtggcag tgcaggccgc cggggccgag 240
gtgaggctgc agctctccgg gcggcggtag cgctggggag gaggaggaga gaag atg 297
Met
1
gcg tcg gag cta gag cca gag gtg cag gcc atc gac cgg agt ttg ctg 345
Ala Ser Glu Leu Glu Pro Glu Val Gln Ala Ile Asp Arg Ser Leu Leu
5 10 15
gaa tgt tcg gcc gag gag att gcg ggg aaa tgg ctg caa gca act gac 393
Glu Cys Ser Ala Glu Glu Ile Ala Gly Lys Trp Leu Gln Ala Thr Asp
20 25 30
ctc act aga gaa gtg tac cag cat tta gcc cac tat gta ccc aaa atc 441
Leu Thr Arg Glu Val Tyr Gln His Leu Ala His Tyr Val Pro Lys Ile
35 40 45
tac tgc agg ggt ccc aac cct ttt cca cag aaa gaa gac atg ctg gca 489
Tyr Cys Arg Gly Pro Asn Pro Phe Pro Gln Lys Glu Asp Met Leu Ala
50 55 60 65
cag cat gtt ttg ttg gga cca atg gaa tgg tac ctt tgt ggt gaa gat 537
Gln His Val Leu Leu Gly Pro Met Glu Trp Tyr Leu Cys Gly Glu Asp
70 75 80
cct gca ttt gga ttt cca aaa ctt gag caa gca aac aaa cct tct cat 585
Pro Ala Phe Gly Phe Pro Lys Leu Glu Gln Ala Asn Lys Pro Ser His
85 90 95

ctt	tgt	ggt	cgt	gtt	ttt	aaa	gta	gga	gag	cct	aca	tat	tct	tgc	aga	633
Leu	Cys	Gly	Arg	Val	Phe	Lys	Val	Gly	Glu	Pro	Thr	Tyr	Ser	Cys	Arg	
	100						105					110				
gac	tgt	gca	gtt	gat	cca	act	tgt	gtt	ttg	tgc	atg	gag	tgc	ttt	ttg	681
Asp	Cys	Ala	Val	Asp	Pro	Thr	Cys	Val	Leu	Cys	Met	Glu	Cys	Phe	Leu	
	115					120					125					
gga	agt	att	cac	aga	gat	cat	cga	tat	agg	atg	aca	aca	tca	gga	ggt	729
Gly	Ser	Ile	His	Arg	Asp	His	Arg	Tyr	Arg	Met	Thr	Thr	Ser	Gly	Gly	
	130				135					140					145	
gga	ggt	ttc	tgt	gac	tgt	ggt	gat	act	gaa	gcc	tggt	aaa	gag	ggt	cct	777
Gly	Gly	Phe	Cys	Asp	Cys	Gly	Asp	Thr	Glu	Ala	Trp	Lys	Glu	Gly	Pro	
			150						155						160	
tac	tgt	caa	aaa	cat	gaa	ctt	aac	acc	tct	gaa	att	gag	gaa	gaa	gag	825
Tyr	Cys	Gln	Lys	His	Glu	Leu	Asn	Thr	Ser	Glu	Ile	Glu	Glu	Glu	Glu	
			165					170						175		
gat	cct	ctt	gtt	cat	tta	tca	gaa	gat	gtg	ata	gca	aga	act	tat	aac	873
Asp	Pro	Leu	Val	His	Leu	Ser	Glu	Asp	Val	Ile	Ala	Arg	Thr	Tyr	Asn	
		180					185					190				
att	ttt	gct	att	acg	ttt	cgg	tat	gca	gta	gaa	ata	tta	acc	tggt	gaa	921
Ile	Phe	Ala	Ile	Thr	Phe	Arg	Tyr	Ala	Val	Glu	Ile	Leu	Thr	Trp	Glu	
	195					200					205					
aaa	gaa	agt	gaa	ttg	cca	gca	gat	tta	gag	atg	gta	gag	aag	agt	gac	969
Lys	Glu	Ser	Glu	Leu	Pro	Ala	Asp	Leu	Glu	Met	Val	Glu	Lys	Ser	Asp	
	210				215					220					225	
acc	tac	tat	tgc	atg	ctg	ttt	aat	gat	gag	gtt	cac	acc	tat	gaa	caa	1017
Thr	Tyr	Tyr	Cys	Met	Leu	Phe	Asn	Asp	Glu	Val	His	Thr	Tyr	Glu	Gln	
			230					235						240		
gtt	att	tat	act	ctt	cag	aaa	gct	gtt	aac	tgt	aca	caa	aaa	gaa	gct	1065
Val	Ile	Tyr	Thr	Leu	Gln	Lys	Ala	Val	Asn	Cys	Thr	Gln	Lys	Glu	Ala	
			245				250						255			
att	ggt	ttt	gca	act	aca	gta	gat	cga	gat	ggg	cgt	agg	tct	gtt	cga	1113
Ile	Gly	Phe	Ala	Thr	Thr	Val	Asp	Arg	Asp	Gly	Arg	Arg	Ser	Val	Arg	
		260					265					270				
tat	gga	gat	ttt	cag	tat	tgt	gag	caa	gca	aaa	tca	gta	att	gtg	aga	1161
Tyr	Gly	Asp	Phe	Gln	Tyr	Cys	Glu	Gln	Ala	Lys	Ser	Val	Ile	Val	Arg	
	275					280					285					
aat	acc	agt	aga	cag	aca	aag	cca	ctc	aaa	gtt	caa	gtt	atg	cat	tcg	1209
Asn	Thr	Ser	Arg	Gln	Thr	Lys	Pro	Leu	Lys	Val	Gln	Val	Met	His	Ser	
	290				295					300					305	
tct	att	gtc	gca	cat	cag	aat	ttt	ggt	ttg	aaa	ctt	ttg	tct	tggt	ctg	1257
Ser	Ile	Val	Ala	His	Gln	Asn	Phe	Gly	Leu	Lys	Leu	Leu	Ser	Trp	Leu	
				310				315						320		
gga	agt	att	att	gga	tat	tca	gat	ggc	ctt	cgc	cgg	att	tta	tgt	caa	1305
Gly	Ser	Ile	Ile	Gly	Tyr	Ser	Asp	Gly	Leu	Arg	Arg	Ile	Leu	Cys	Gln	
			325					330					335			
gtt	ggt	tta	caa	gaa	ggg	cca	gat	ggt	gaa	aac	tct	tct	cta	gtg	gac	1353
Val	Gly	Leu	Gln	Glu	Gly	Pro	Asp	Gly	Glu	Asn	Ser	Ser	Leu	Val	Asp	
		340					345					350				

aga ctg atg ctt agt gat tcc aaa tta tgg aaa ggt gct agg agt gta	1401
Arg Leu Met Leu Ser Asp Ser Lys Leu Trp Lys Gly Ala Arg Ser Val	
355 360 365	
tat cat cag ttg ttc atg agc agt ctg ctt atg gat ttg aaa tac aag	1449
Tyr His Gln Leu Phe Met Ser Ser Leu Leu Met Asp Leu Lys Tyr Lys	
370 375 380 385	
aaa cta ttt gct gtt cga ttt gca aaa aat tac cag cag ttg cag aga	1497
Lys Leu Phe Ala Val Arg Phe Ala Lys Asn Tyr Gln Gln Leu Gln Arg	
390 395 400	
gat ttt atg gag gat gat cac gag cga gca gtg tcg gtg act gct cta	1545
Asp Phe Met Glu Asp Asp His Glu Arg Ala Val Ser Val Thr Ala Leu	
405 410 415	
tct gtc cag ttc ttc acc gca cct act ctg gct cga atg ctc atc aca	1593
Ser Val Gln Phe Phe Thr Ala Pro Thr Leu Ala Arg Met Leu Ile Thr	
420 425 430	
gaa gaa aac ttg atg agc att atc att aag act ttt atg gat cat ttg	1641
Glu Glu Asn Leu Met Ser Ile Ile Ile Lys Thr Phe Met Asp His Leu	
435 440 445	
aga cat cga gat gcc cag ggc aga ttt cag ttt gaa cga tac act gct	1689
Arg His Arg Asp Ala Gln Gly Arg Phe Gln Phe Glu Arg Tyr Thr Ala	
450 455 460 465	
tta caa gcc ttc aaa ttt agg aga gta cag agc ctt att tta gat ctc	1737
Leu Gln Ala Phe Lys Phe Arg Arg Val Gln Ser Leu Ile Leu Asp Leu	
470 475 480	
aag tat gtg tta att agc aaa cca act gaa tgg tca gat gag ctg agg	1785
Lys Tyr Val Leu Ile Ser Lys Pro Thr Glu Trp Ser Asp Glu Leu Arg	
485 490 495	
cag aag ttc cta gaa ggg ttt gat gcc ttt ttg gaa tta cta aaa tgt	1833
Gln Lys Phe Leu Glu Gly Phe Asp Ala Phe Leu Glu Leu Leu Lys Cys	
500 505 510	
atg cag gga atg gat cca att aca cgt caa gta gga caa cat att gaa	1881
Met Gln Gly Met Asp Pro Ile Thr Arg Gln Val Gly Gln His Ile Glu	
515 520 525	
atg gaa cca gag tgg gaa gca gcc ttc aca cta caa atg aaa tta aca	1929
Met Glu Pro Glu Trp Glu Ala Ala Phe Thr Leu Gln Met Lys Leu Thr	
530 535 540 545	
cat gtc att tca atg atg cag gac tgg tgt gct tca gat gaa aaa gtg	1977
His Val Ile Ser Met Met Gln Asp Trp Cys Ala Ser Asp Glu Lys Val	
550 555 560	
tta atc gaa gct tac aag aaa tgt ctc gct gta ctg atg cag tgt cat	2025
Leu Ile Glu Ala Tyr Lys Lys Cys Leu Ala Val Leu Met Gln Cys His	
565 570 575	
ggg ggt tat act gat ggt gaa cag cca atc aca cta agc att tgt gga	2073
Gly Gly Tyr Thr Asp Gly Glu Gln Pro Ile Thr Leu Ser Ile Cys Gly	
580 585 590	
cat tca gtg gaa act atc aga tac tgt gtt tcc caa gaa aaa gtt agc	2121
His Ser Val Glu Thr Ile Arg Tyr Cys Val Ser Gln Glu Lys Val Ser	
595 600 605	

GenBank accession number: U00096.2 (H. pylori strain NCTC 11637 genome)

att cac ctc cca gtt tct cgc tta ctt gca ggt tta cat gta tta tta	2169
Ile His Leu Pro Val Ser Arg Leu Leu Ala Gly Leu His Val Leu Leu	
610 615 620 625	
agc aaa agt gaa gtg gca tat aaa ttt cca gag ctc cta cct cta agt	2217
Ser Lys Ser Glu Val Ala Tyr Lys Phe Pro Glu Leu Leu Pro Leu Ser	
630 635 640	
gaa ctt agc cca ccc atg ttg ata gaa cac cct ctt aga tgt ctt gtt	2265
Glu Leu Ser Pro Pro Met Leu Ile Glu His Pro Leu Arg Cys Leu Val	
645 650 655	
ctg tgt gcc caa gta cat gcc gga atg tgg aga aga aat ggg ttc tct	2313
Leu Cys Ala Gln Val His Ala Gly Met Trp Arg Arg Asn Gly Phe Ser	
660 665 670	
cta gta aac cag att tat tac tac cat aat gtg aaa tgc aga cgt gag	2361
Leu Val Asn Gln Ile Tyr Tyr Tyr His Asn Val Lys Cys Arg Arg Glu	
675 680 685	
atg ttt gac aag gat gta gta atg ctt cag aca ggt gtc tcc atg atg	2409
Met Phe Asp Lys Asp Val Val Met Leu Gln Thr Gly Val Ser Met Met	
690 695 700 705	
gat cca aat cat ttc ctg atg atc atg ctc agc cgc ttt gaa ctt tat	2457
Asp Pro Asn His Phe Leu Met Ile Met Leu Ser Arg Phe Glu Leu Tyr	
710 715 720	
cag att ttc agt act cca gac tat gga aaa aga ttt agt tct gag att	2505
Gln Ile Phe Ser Thr Pro Asp Tyr Gly Lys Arg Phe Ser Ser Glu Ile	
725 730 735	
acc cat aag gat gtt gtt cag cag aac aat act cta ata gaa gaa atg	2553
Thr His Lys Asp Val Val Gln Gln Asn Asn Thr Leu Ile Glu Glu Met	
740 745 750	
cta tac ctc att ata atg ctt gtt gga gag aga ttt agt cct gga gtt	2601
Leu Tyr Leu Ile Ile Met Leu Val Gly Glu Arg Phe Ser Pro Gly Val	
755 760 765	
gga cag gta aat gct aca gat gaa atc aag cga gag att atc cat cag	2649
Gly Gln Val Asn Ala Thr Asp Glu Ile Lys Arg Glu Ile Ile His Gln	
770 775 780 785	
ttg agt atc aag cct atg gct cat agt gaa ttg gta aag tct tta cct	2697
Leu Ser Ile Lys Pro Met Ala His Ser Glu Leu Val Lys Ser Leu Pro	
790 795 800	
gaa gat gag aac aag gag act ggc atg gag agt gta atc gaa gca gtt	2745
Glu Asp Glu Asn Lys Glu Thr Gly Met Glu Ser Val Ile Glu Ala Val	
805 810 815	
gcc cat ttc aag aaa cct gga tta aca gga cga ggc atg tat gaa ctg	2793
Ala His Phe Lys Lys Pro Gly Leu Thr Gly Arg Gly Met Tyr Glu Leu	
820 825 830	
aaa cca gaa tgt gcc aaa gag ttc aac ttg tat ttc tat cac ttt tca	2841
Lys Pro Glu Cys Ala Lys Glu Phe Asn Leu Tyr Phe Tyr His Phe Ser	
835 840 845	
agg gca gaa cag tcc aag gca gaa gaa gcg caa cgg aaa ttg aaa aga	2889
Arg Ala Glu Gln Ser Lys Ala Glu Glu Ala Gln Arg Lys Leu Lys Arg	
850 855 860 865	

caa aat aga gaa gat aca gca ctc cca cct ccg gtg ttg cct cca ttc	2937
Gln Asn Arg Glu Asp Thr Ala Leu Pro Pro Pro Val Leu Pro Pro Phe	
870 875 880	
tgc cct ctg ttt gca agc ctg gtt aac att ttg cag tca gat gtc atg	2985
Cys Pro Leu Phe Ala Ser Leu Val Asn Ile Leu Gln Ser Asp Val Met	
885 890 895	
ttg tgc atc atg gga aca att ctg caa tgg gct gtg gaa cat aat gga	3033
Leu Cys Ile Met Gly Thr Ile Leu Gln Trp Ala Val Glu His Asn Gly	
900 905 910	
tat gcc tgg tca gag tcc atg ctg caa agg gtg tta cat tta att ggc	3081
Tyr Ala Trp Ser Glu Ser Met Leu Gln Arg Val Leu His Leu Ile Gly	
915 920 925	
atg gca cta caa gaa gaa aaa caa cat tta gag aat gtc acg gaa gag	3129
Met Ala Leu Gln Glu Glu Lys Gln His Leu Glu Asn Val Thr Glu Glu	
930 935 940 945	
cat gta gta aca ttt acc ttc act cag aag ata tca aaa cct ggt gaa	3177
His Val Val Thr Phe Thr Phe Thr Gln Lys Ile Ser Lys Pro Gly Glu	
950 955 960	
gcg cca aaa aat tct cct agc ata cta gct atg ctg gaa aca cta caa	3225
Ala Pro Lys Asn Ser Pro Ser Ile Leu Ala Met Leu Glu Thr Leu Gln	
965 970 975	
aat gct ccc tac cta gaa gtc cac aaa gac atg att cgg tgg ata ttg	3273
Asn Ala Pro Tyr Leu Glu Val His Lys Asp Met Ile Arg Trp Ile Leu	
980 985 990	
aag act ttt aat gct gtt aaa aag atg agg gag agt tca cct acc agt	3321
Lys Thr Phe Asn Ala Val Lys Lys Met Arg Glu Ser Ser Pro Thr Ser	
995 1000 1005	
ccc gtg gca gag aca gaa gga acc ata atg gaa gag agt tca agg gac	3369
Pro Val Ala Glu Thr Glu Gly Thr Ile Met Glu Glu Ser Ser Arg Asp	
1010 1015 1020 1025	
aaa gac aaa gct gag agg aag aga aaa gca gag att gcc aga ctg cgc	3417
Lys Asp Lys Ala Glu Arg Lys Arg Lys Ala Glu Ile Ala Arg Leu Arg	
1030 1035 1040	
aga gaa aag atc atg gct cag atg tct gaa atg cag cgg cat ttt att	3465
Arg Glu Lys Ile Met Ala Gln Met Ser Glu Met Gln Arg His Phe Ile	
1045 1050 1055	
gat gaa aac aaa gaa ctc ttt cag cag aca tta gaa ctg gat gcc tca	3513
Asp Glu Asn Lys Glu Leu Phe Gln Gln Thr Leu Glu Leu Asp Ala Ser	
1060 1065 1070	
acc tct gct gtt ctt gat cat agc cct gtg gct tca gat atg aca ctt	3561
Thr Ser Ala Val Leu Asp His Ser Pro Val Ala Ser Asp Met Thr Leu	
1075 1080 1085	
aca gca ctg ggt ccc aca caa act cag gtt cct gaa caa aga caa ttc	3609
Thr Ala Leu Gly Pro Thr Gln Thr Gln Val Pro Glu Gln Arg Gln Phe	
1090 1095 1100 1105	
gtt aca tgt ata ttg tgt caa gag gag caa gaa gtt aaa gtg gaa agc	3657
Val Thr Cys Ile Leu Cys Gln Glu Glu Gln Glu Val Lys Val Glu Ser	
1110 1115 1120	

agg gca atg gtc ttg gca gca ttt gtt cag aga tca act gta tta tca	3705
Arg Ala Met Val Leu Ala Ala Phe Val Gln Arg Ser Thr Val Leu Ser	
1125 1130 1135	
aaa aac aga agt aaa ttt att caa gat cca gaa aaa tat gat cca tta	3753
Lys Asn Arg Ser Lys Phe Ile Gln Asp Pro Glu Lys Tyr Asp Pro Leu	
1140 1145 1150	
ttc atg cac cct gat ctg tct tgt gga aca cac act agt agc tgt ggg	3801
Phe Met His Pro Asp Leu Ser Cys Gly Thr His Thr Ser Ser Cys Gly	
1155 1160 1165	
cac att atg cat gcc cat tgt tgg caa agg tat ttt gat tcc gtt caa	3849
His Ile Met His Ala His Cys Trp Gln Arg Tyr Phe Asp Ser Val Gln	
1170 1175 1180 1185	
gct aaa gaa cag cga agg caa cag aga tta cgc tta cat acg agc tat	3897
Ala Lys Glu Gln Arg Arg Gln Gln Arg Leu Arg Leu His Thr Ser Tyr	
1190 1195 1200	
gat gta gaa aac gga gaa ttc ctt tgc ccc ctt tgt gaa tgc ttg agt	3945
Asp Val Glu Asn Gly Glu Phe Leu Cys Pro Leu Cys Glu Cys Leu Ser	
1205 1210 1215	
aat act gtt att cct ctg ctg ctt cct cca aga aat att ttt aac aac	3993
Asn Thr Val Ile Pro Leu Leu Leu Pro Pro Arg Asn Ile Phe Asn Asn	
1220 1225 1230	
agg tta aat ttt tca gac caa cca aat ctg act cag tgg att aga aca	4041
Arg Leu Asn Phe Ser Asp Gln Pro Asn Leu Thr Gln Trp Ile Arg Thr	
1235 1240 1245	
ata tct cag caa ata aaa gca tta cag ttt ctt agg aaa gaa gaa agt	4089
Ile Ser Gln Gln Ile Lys Ala Leu Gln Phe Leu Arg Lys Glu Glu Ser	
1250 1255 1260 1265	
act cct aat aat gcc tct aca aag aat tca gaa aat gtg gat gaa tta	4137
Thr Pro Asn Asn Ala Ser Thr Lys Asn Ser Glu Asn Val Asp Glu Leu	
1270 1275 1280	
cag ctc cct gaa ggg ttc agg cct gat ttt cgt cct aag atc cct tat	4185
Gln Leu Pro Glu Gly Phe Arg Pro Asp Phe Arg Pro Lys Ile Pro Tyr	
1285 1290 1295	
tct gag agc ata aaa gaa atg cta acg aca ttt gga act gct acc tac	4233
Ser Glu Ser Ile Lys Glu Met Leu Thr Thr Phe Gly Thr Ala Thr Tyr	
1300 1305 1310	
aag gtg gga cta aag gtt cat ccc aat gaa gag gat cct cgt gtt ccc	4281
Lys Val Gly Leu Lys Val His Pro Asn Glu Glu Asp Pro Arg Val Pro	
1315 1320 1325	
ata atg tgt tgg ggt agc tgc gcg tac acc atc caa agc ata gaa aga	4329
Ile Met Cys Trp Gly Ser Cys Ala Tyr Thr Ile Gln Ser Ile Glu Arg	
1330 1335 1340 1345	
att ttg agt gat gaa gat aaa cca ttg ttt ggt cct tta cct tgc aga	4377
Ile Leu Ser Asp Glu Asp Lys Pro Leu Phe Gly Pro Leu Pro Cys Arg	
1350 1355 1360	
ctg gat gac tgt ctt agg tca ttg acg aga ttt gcc gca gca cac tgg	4425
Leu Asp Asp Cys Leu Arg Ser Leu Thr Arg Phe Ala Ala Ala His Trp	
1365 1370 1375	

123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

aca gtg gca tca gtt tca gtg gtg caa gga cat ttt tgt aaa ctt ttt	4473
Thr Val Ala Ser Val Ser Val Val Gln Gly His Phe Cys Lys Leu Phe	
1380 1385 1390	
gca tca ctg gtg cct aat gac agc cat gag gaa ctt cca tgc ata tta	4521
Ala Ser Leu Val Pro Asn Asp Ser His Glu Glu Leu Pro Cys Ile Leu	
1395 1400 1405	
gat att gac atg ttt cat tta ttg gtg ggc ttg gtg ctt gca ttt cct	4569
Asp Ile Asp Met Phe His Leu Leu Val Gly Leu Val Leu Ala Phe Pro	
1410 1415 1420 1425	
gcg ttg cag tgt cag gat ttt tca ggg atc agc ctt ggc act gga gac	4617
Ala Leu Gln Cys Gln Asp Phe Ser Gly Ile Ser Leu Gly Thr Gly Asp	
1430 1435 1440	
ctt cac att ttc cat ctg gtt act atg gca cac atc ata cag atc tta	4665
Leu His Ile Phe His Leu Val Thr Met Ala His Ile Ile Gln Ile Leu	
1445 1450 1455	
ctt acc tca tgt aca gaa gag aat ggc atg gat caa gaa aat ccc cct	4713
Leu Thr Ser Cys Thr Glu Glu Asn Gly Met Asp Gln Glu Asn Pro Pro	
1460 1465 1470	
tgt gaa gaa gaa tca gca gtt ctt gct ttg tat aaa aca ctt cac cag	4761
Cys Glu Glu Glu Ser Ala Val Leu Ala Leu Tyr Lys Thr Leu His Gln	
1475 1480 1485	
tat acg gga agt gcc ttg aaa gaa ata cca tcc ggc tgg cat ctg tgg	4809
Tyr Thr Gly Ser Ala Leu Lys Glu Ile Pro Ser Gly Trp His Leu Trp	
1490 1495 1500 1505	
agg agt gtc aga gct gga atc atg cct ttc ctg aag tgt tct gct tta	4857
Arg Ser Val Arg Ala Gly Ile Met Pro Phe Leu Lys Cys Ser Ala Leu	
1510 1515 1520	
ttt ttt cat tac tta aat gga gtt cct tcc cca ccc gac att caa gtt	4905
Phe Phe His Tyr Leu Asn Gly Val Pro Ser Pro Pro Asp Ile Gln Val	
1525 1530 1535	
cct gga aca agc cat ttt gaa cat tta tgt agc tat ctt tcc cta cca	4953
Pro Gly Thr Ser His Phe Glu His Leu Cys Ser Tyr Leu Ser Leu Pro	
1540 1545 1550	
aac aac ctc att tgc ctt ttt caa gaa aat agt gag ata atg aat tca	5001
Asn Asn Leu Ile Cys Leu Phe Gln Glu Asn Ser Glu Ile Met Asn Ser	
1555 1560 1565	
ctg att gaa agt tgg tgc cgt aac agt gaa gtt aaa aga tat cta gaa	5049
Leu Ile Glu Ser Trp Cys Arg Asn Ser Glu Val Lys Arg Tyr Leu Glu	
1570 1575 1580 1585	
ggg gaa aga gat gct ata aga tat cca aga gaa tct aac aaa tta ata	5097
Gly Glu Arg Asp Ala Ile Arg Tyr Pro Arg Glu Ser Asn Lys Leu Ile	
1590 1595 1600	
aac ctt cca gag gat tac agc agc ctc att aat caa gca tcc aat ttc	5145
Asn Leu Pro Glu Asp Tyr Ser Ser Leu Ile Asn Gln Ala Ser Asn Phe	
1605 1610 1615	
tcg tgc ccg aaa tca ggt ggt gat aag agc aga gcc cca act ctg tgc	5193
Ser Cys Pro Lys Ser Gly Gly Asp Lys Ser Arg Ala Pro Thr Leu Cys	
1620 1625 1630	

ctt gtg tgc gga tct ctg ctg tgc tcc cag agt tac tgc tgc cag act 5241
 Leu Val Cys Gly Ser Leu Leu Cys Ser Gln Ser Tyr Cys Cys Gln Thr
 1635 1640 1645
 gaa ctg gaa ggg gag gat gta gga gcc tgc aca gct cac acc tac tcc 5289
 Glu Leu Glu Gly Glu Asp Val Gly Ala Cys Thr Ala His Thr Tyr Ser
 1650 1655 1660 1665
 tgt ggc tct gga gtg ggc atc ttc ctg aga gta cgg gaa tgt cag gtg 5337
 Cys Gly Ser Gly Val Gly Ile Phe Leu Arg Val Arg Glu Cys Gln Val
 1670 1675 1680
 cta ttt tta gct ggc aaa acc aaa ggc tgt ttt tat tct cct cct tac 5385
 Leu Phe Leu Ala Gly Lys Thr Lys Gly Cys Phe Tyr Ser Pro Pro Tyr
 1685 1690 1695
 ctt gat gac tat ggg gag acc gac cag gga ctc aga cgg gga aat cct 5433
 Leu Asp Asp Tyr Gly Glu Thr Asp Gln Gly Leu Arg Arg Gly Asn Pro
 1700 1705 1710
 tta cat tta tgc aaa gag cga ttc aag aag att cag aag ctc tgg cac 5481
 Leu His Leu Cys Lys Glu Arg Phe Lys Lys Ile Gln Lys Leu Trp His
 1715 1720 1725
 caa cac agt gtc aca gag gaa att gga cat gca cag gaa gcc aat cag 5529
 Gln His Ser Val Thr Glu Glu Ile Gly His Ala Gln Glu Ala Asn Gln
 1730 1735 1740 1745
 aca ctg gtt ggc att gac tgg caa cat tta taattattgc accacaaaaa 5579
 Thr Leu Val Gly Ile Asp Trp Gln His Leu
 1750 1755
 aacacaaaact tggatttttt taaccagtt ggctttttta gaaagaaaga agttctgctg 5639
 aatttggaaa taaattcttt atttaaactt tccttcccag ttttatagtt tctgggttctg 5699
 aggactgatg aaaatcatct tccatcagca gattttcttg cactgtttgc tgtgcccctc 5759
 aaatataatg tcttggggtt taagatcgag caaggagctt ctcttcctag attggatccc 5819
 agcccctttg tgggggtctg actgcatagt ccagccatt atgtgatatt tcacgttatt 5879
 gatgatagt aaccgtgggt ccgaagctga ctcaacggag gcaggaaca agtctctgt 5939
 ggtctgttg gtcatacttc ctggttccac tgagtggccc aacctggga ctgggttgg 5999
 gtcccctctg ctgacaggac cctactccta ggagcaaagt ggttgatttt gaaggcagt 6059
 ttcccttctc tccattgact atgagagagt tgggggacac acatgcagaa gaagcccgtg 6119
 gggagaagggt ggattcctgg tgtgctggct gggttttcag ggctgtaga ggtttttttt 6179
 ttcttttttt ttttatggc aagacttttg gctttgagaa aactcactta gagggctttc 6239
 caaaaactta ggatgggtcta aaaaattagg atattctttt agaattagga agaaaaatta 6299
 g 6300

<210> 4
 <211> 1755
 <212> PRT
 <213> Homo sapiens

<400> 4

Met	Ala	Ser	Glu	Leu	Glu	Pro	Glu	Val	Gln	Ala	Ile	Asp	Arg	Ser	Leu	1	5	10	15
Leu	Glu	Cys	Ser	Ala	Glu	Glu	Ile	Ala	Gly	Lys	Trp	Leu	Gln	Ala	Thr	20	25	30	
Asp	Leu	Thr	Arg	Glu	Val	Tyr	Gln	His	Leu	Ala	His	Tyr	Val	Pro	Lys	35	40	45	
Ile	Tyr	Cys	Arg	Gly	Pro	Asn	Pro	Phe	Pro	Gln	Lys	Glu	Asp	Met	Leu	50	55	60	
Ala	Gln	His	Val	Leu	Leu	Gly	Pro	Met	Glu	Trp	Tyr	Leu	Cys	Gly	Glu	65	70	75	80
Asp	Pro	Ala	Phe	Gly	Phe	Pro	Lys	Leu	Glu	Gln	Ala	Asn	Lys	Pro	Ser	85	90	95	
His	Leu	Cys	Gly	Arg	Val	Phe	Lys	Val	Gly	Glu	Pro	Thr	Tyr	Ser	Cys	100	105	110	
Arg	Asp	Cys	Ala	Val	Asp	Pro	Thr	Cys	Val	Leu	Cys	Met	Glu	Cys	Phe	115	120	125	
Leu	Gly	Ser	Ile	His	Arg	Asp	His	Arg	Tyr	Arg	Met	Thr	Thr	Ser	Gly	130	135	140	
Gly	Gly	Gly	Phe	Cys	Asp	Cys	Gly	Asp	Thr	Glu	Ala	Trp	Lys	Glu	Gly	145	150	155	160
Pro	Tyr	Cys	Gln	Lys	His	Glu	Leu	Asn	Thr	Ser	Glu	Ile	Glu	Glu	Glu	165	170	175	
Glu	Asp	Pro	Leu	Val	His	Leu	Ser	Glu	Asp	Val	Ile	Ala	Arg	Thr	Tyr	180	185	190	
Asn	Ile	Phe	Ala	Ile	Thr	Phe	Arg	Tyr	Ala	Val	Glu	Ile	Leu	Thr	Trp	195	200	205	
Glu	Lys	Glu	Ser	Glu	Leu	Pro	Ala	Asp	Leu	Glu	Met	Val	Glu	Lys	Ser	210	215	220	
Asp	Thr	Tyr	Tyr	Cys	Met	Leu	Phe	Asn	Asp	Glu	Val	His	Thr	Tyr	Glu	225	230	235	240
Gln	Val	Ile	Tyr	Thr	Leu	Gln	Lys	Ala	Val	Asn	Cys	Thr	Gln	Lys	Glu	245	250	255	
Ala	Ile	Gly	Phe	Ala	Thr	Thr	Val	Asp	Arg	Asp	Gly	Arg	Arg	Ser	Val	260	265	270	
Arg	Tyr	Gly	Asp	Phe	Gln	Tyr	Cys	Glu	Gln	Ala	Lys	Ser	Val	Ile	Val	275	280	285	
Arg	Asn	Thr	Ser	Arg	Gln	Thr	Lys	Pro	Leu	Lys	Val	Gln	Val	Met	His	290	295	300	
Ser	Ser	Ile	Val	Ala	His	Gln	Asn	Phe	Gly	Leu	Lys	Leu	Leu	Ser	Trp	305	310	315	320
Leu	Gly	Ser	Ile	Ile	Gly	Tyr	Ser	Asp	Gly	Leu	Arg	Arg	Ile	Leu	Cys	325	330	335	

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400

Gln	Val	Gly	Leu	Gln	Glu	Gly	Pro	Asp	Gly	Glu	Asn	Ser	Ser	Leu	Val
			340					345					350		
Asp	Arg	Leu	Met	Leu	Ser	Asp	Ser	Lys	Leu	Trp	Lys	Gly	Ala	Arg	Ser
		355				360						365			
Val	Tyr	His	Gln	Leu	Phe	Met	Ser	Ser	Leu	Leu	Met	Asp	Leu	Lys	Tyr
	370					375					380				
Lys	Lys	Leu	Phe	Ala	Val	Arg	Phe	Ala	Lys	Asn	Tyr	Gln	Gln	Leu	Gln
385					390					395					400
Arg	Asp	Phe	Met	Glu	Asp	Asp	His	Glu	Arg	Ala	Val	Ser	Val	Thr	Ala
				405					410					415	
Leu	Ser	Val	Gln	Phe	Phe	Thr	Ala	Pro	Thr	Leu	Ala	Arg	Met	Leu	Ile
			420					425					430		
Thr	Glu	Glu	Asn	Leu	Met	Ser	Ile	Ile	Ile	Lys	Thr	Phe	Met	Asp	His
		435				440						445			
Leu	Arg	His	Arg	Asp	Ala	Gln	Gly	Arg	Phe	Gln	Phe	Glu	Arg	Tyr	Thr
	450					455					460				
Ala	Leu	Gln	Ala	Phe	Lys	Phe	Arg	Arg	Val	Gln	Ser	Leu	Ile	Leu	Asp
465					470					475					480
Leu	Lys	Tyr	Val	Leu	Ile	Ser	Lys	Pro	Thr	Glu	Trp	Ser	Asp	Glu	Leu
				485					490					495	
Arg	Gln	Lys	Phe	Leu	Glu	Gly	Phe	Asp	Ala	Phe	Leu	Glu	Leu	Leu	Lys
			500					505					510		
Cys	Met	Gln	Gly	Met	Asp	Pro	Ile	Thr	Arg	Gln	Val	Gly	Gln	His	Ile
		515					520					525			
Glu	Met	Glu	Pro	Glu	Trp	Glu	Ala	Ala	Phe	Thr	Leu	Gln	Met	Lys	Leu
	530					535					540				
Thr	His	Val	Ile	Ser	Met	Met	Gln	Asp	Trp	Cys	Ala	Ser	Asp	Glu	Lys
545					550					555					560
Val	Leu	Ile	Glu	Ala	Tyr	Lys	Lys	Cys	Leu	Ala	Val	Leu	Met	Gln	Cys
				565					570					575	
His	Gly	Gly	Tyr	Thr	Asp	Gly	Glu	Gln	Pro	Ile	Thr	Leu	Ser	Ile	Cys
			580					585					590		
Gly	His	Ser	Val	Glu	Thr	Ile	Arg	Tyr	Cys	Val	Ser	Gln	Glu	Lys	Val
		595					600					605			
Ser	Ile	His	Leu	Pro	Val	Ser	Arg	Leu	Leu	Ala	Gly	Leu	His	Val	Leu
	610					615					620				
Leu	Ser	Lys	Ser	Glu	Val	Ala	Tyr	Lys	Phe	Pro	Glu	Leu	Leu	Pro	Leu
625					630					635					640
Ser	Glu	Leu	Ser	Pro	Pro	Met	Leu	Ile	Glu	His	Pro	Leu	Arg	Cys	Leu
				645					650					655	
Val	Leu	Cys	Ala	Gln	Val	His	Ala	Gly	Met	Trp	Arg	Arg	Asn	Gly	Phe
			660					665					670		

Ser	Leu	Val	Asn	Gln	Ile	Tyr	Tyr	His	Asn	Val	Lys	Cys	Arg	Arg	
		675					680				685				
Glu	Met	Phe	Asp	Lys	Asp	Val	Val	Met	Leu	Gln	Thr	Gly	Val	Ser	Met
	690					695					700				
Met	Asp	Pro	Asn	His	Phe	Leu	Met	Ile	Met	Leu	Ser	Arg	Phe	Glu	Leu
705					710					715					720
Tyr	Gln	Ile	Phe	Ser	Thr	Pro	Asp	Tyr	Gly	Lys	Arg	Phe	Ser	Ser	Glu
				725					730					735	
Ile	Thr	His	Lys	Asp	Val	Val	Gln	Gln	Asn	Asn	Thr	Leu	Ile	Glu	Glu
			740					745					750		
Met	Leu	Tyr	Leu	Ile	Ile	Met	Leu	Val	Gly	Glu	Arg	Phe	Ser	Pro	Gly
		755					760					765			
Val	Gly	Gln	Val	Asn	Ala	Thr	Asp	Glu	Ile	Lys	Arg	Glu	Ile	Ile	His
	770					775					780				
Gln	Leu	Ser	Ile	Lys	Pro	Met	Ala	His	Ser	Glu	Leu	Val	Lys	Ser	Leu
785					790					795					800
Pro	Glu	Asp	Glu	Asn	Lys	Glu	Thr	Gly	Met	Glu	Ser	Val	Ile	Glu	Ala
				805					810					815	
Val	Ala	His	Phe	Lys	Lys	Pro	Gly	Leu	Thr	Gly	Arg	Gly	Met	Tyr	Glu
			820					825					830		
Leu	Lys	Pro	Glu	Cys	Ala	Lys	Glu	Phe	Asn	Leu	Tyr	Phe	Tyr	His	Phe
		835					840					845			
Ser	Arg	Ala	Glu	Gln	Ser	Lys	Ala	Glu	Glu	Ala	Gln	Arg	Lys	Leu	Lys
	850					855					860				
Arg	Gln	Asn	Arg	Glu	Asp	Thr	Ala	Leu	Pro	Pro	Pro	Val	Leu	Pro	Pro
865					870					875					880
Phe	Cys	Pro	Leu	Phe	Ala	Ser	Leu	Val	Asn	Ile	Leu	Gln	Ser	Asp	Val
				885					890					895	
Met	Leu	Cys	Ile	Met	Gly	Thr	Ile	Leu	Gln	Trp	Ala	Val	Glu	His	Asn
			900					905					910		
Gly	Tyr	Ala	Trp	Ser	Glu	Ser	Met	Leu	Gln	Arg	Val	Leu	His	Leu	Ile
		915					920					925			
Gly	Met	Ala	Leu	Gln	Glu	Glu	Lys	Gln	His	Leu	Glu	Asn	Val	Thr	Glu
	930					935					940				
Glu	His	Val	Val	Thr	Phe	Thr	Phe	Thr	Gln	Lys	Ile	Ser	Lys	Pro	Gly
945					950					955					960
Glu	Ala	Pro	Lys	Asn	Ser	Pro	Ser	Ile	Leu	Ala	Met	Leu	Glu	Thr	Leu
				965					970					975	
Gln	Asn	Ala	Pro	Tyr	Leu	Glu	Val	His	Lys	Asp	Met	Ile	Arg	Trp	Ile
			980					985					990		
Leu	Lys	Thr	Phe	Asn	Ala	Val	Lys	Lys	Met	Arg	Glu	Ser	Ser	Pro	Thr
		995					1000					1005			

Ser	Pro	Val	Ala	Glu	Thr	Glu	Gly	Thr	Ile	Met	Glu	Glu	Ser	Ser	Arg
1010						1015				1020					
Asp	Lys	Asp	Lys	Ala	Glu	Arg	Lys	Arg	Lys	Ala	Glu	Ile	Ala	Arg	Leu
025					1030				1035						1040
Arg	Arg	Glu	Lys	Ile	Met	Ala	Gln	Met	Ser	Glu	Met	Gln	Arg	His	Phe
				1045				1050						1055	
Ile	Asp	Glu	Asn	Lys	Glu	Leu	Phe	Gln	Gln	Thr	Leu	Glu	Leu	Asp	Ala
		1060						1065						1070	
Ser	Thr	Ser	Ala	Val	Leu	Asp	His	Ser	Pro	Val	Ala	Ser	Asp	Met	Thr
1075						1080						1085			
Leu	Thr	Ala	Leu	Gly	Pro	Thr	Gln	Thr	Gln	Val	Pro	Glu	Gln	Arg	Gln
1090						1095						1100			
Phe	Val	Thr	Cys	Ile	Leu	Cys	Gln	Glu	Glu	Gln	Glu	Val	Lys	Val	Glu
105					1110				1115						1120
Ser	Arg	Ala	Met	Val	Leu	Ala	Ala	Phe	Val	Gln	Arg	Ser	Thr	Val	Leu
				1125				1130						1135	
Ser	Lys	Asn	Arg	Ser	Lys	Phe	Ile	Gln	Asp	Pro	Glu	Lys	Tyr	Asp	Pro
		1140						1145				1150			
Leu	Phe	Met	His	Pro	Asp	Leu	Ser	Cys	Gly	Thr	His	Thr	Ser	Ser	Cys
1155						1160						1165			
Gly	His	Ile	Met	His	Ala	His	Cys	Trp	Gln	Arg	Tyr	Phe	Asp	Ser	Val
1170						1175				1180					
Gln	Ala	Lys	Glu	Gln	Arg	Arg	Gln	Gln	Arg	Leu	Arg	Leu	His	Thr	Ser
185					1190				1195						1200
Tyr	Asp	Val	Glu	Asn	Gly	Glu	Phe	Leu	Cys	Pro	Leu	Cys	Glu	Cys	Leu
		1205						1210				1215			
Ser	Asn	Thr	Val	Ile	Pro	Leu	Leu	Leu	Pro	Pro	Arg	Asn	Ile	Phe	Asn
		1220						1225				1230			
Asn	Arg	Leu	Asn	Phe	Ser	Asp	Gln	Pro	Asn	Leu	Thr	Gln	Trp	Ile	Arg
1235						1240						1245			
Thr	Ile	Ser	Gln	Gln	Ile	Lys	Ala	Leu	Gln	Phe	Leu	Arg	Lys	Glu	Glu
1250						1255				1260					
Ser	Thr	Pro	Asn	Asn	Ala	Ser	Thr	Lys	Asn	Ser	Glu	Asn	Val	Asp	Glu
265					1270				1275						1280
Leu	Gln	Leu	Pro	Glu	Gly	Phe	Arg	Pro	Asp	Phe	Arg	Pro	Lys	Ile	Pro
		1285						1290				1295			
Tyr	Ser	Glu	Ser	Ile	Lys	Glu	Met	Leu	Thr	Thr	Phe	Gly	Thr	Ala	Thr
		1300						1305				1310			
Tyr	Lys	Val	Gly	Leu	Lys	Val	His	Pro	Asn	Glu	Glu	Asp	Pro	Arg	Val
1315						1320						1325			
Pro	Ile	Met	Cys	Trp	Gly	Ser	Cys	Ala	Tyr	Thr	Ile	Gln	Ser	Ile	Glu
1330						1335				1340					

Arg	Ile	Leu	Ser	Asp	Glu	Asp	Lys	Pro	Leu	Phe	Gly	Pro	Leu	Pro	Cys	
345					1350					1355					1360	
Arg	Leu	Asp	Asp	Cys	Leu	Arg	Ser	Leu	Thr	Arg	Phe	Ala	Ala	Ala	His	
				1365					1370						1375	
Trp	Thr	Val	Ala	Ser	Val	Ser	Val	Val	Gln	Gly	His	Phe	Cys	Lys	Leu	
			1380					1385					1390			
Phe	Ala	Ser	Leu	Val	Pro	Asn	Asp	Ser	His	Glu	Glu	Leu	Pro	Cys	Ile	
		1395					1400					1405				
Leu	Asp	Ile	Asp	Met	Phe	His	Leu	Leu	Val	Gly	Leu	Val	Leu	Ala	Phe	
	1410					1415					1420					
Pro	Ala	Leu	Gln	Cys	Gln	Asp	Phe	Ser	Gly	Ile	Ser	Leu	Gly	Thr	Gly	
425					1430					1435					1440	
Asp	Leu	His	Ile	Phe	His	Leu	Val	Thr	Met	Ala	His	Ile	Ile	Gln	Ile	
				1445					1450						1455	
Leu	Leu	Thr	Ser	Cys	Thr	Glu	Glu	Asn	Gly	Met	Asp	Gln	Glu	Asn	Pro	
		1460						1465						1470		
Pro	Cys	Glu	Glu	Glu	Ser	Ala	Val	Leu	Ala	Leu	Tyr	Lys	Thr	Leu	His	
		1475					1480					1485				
Gln	Tyr	Thr	Gly	Ser	Ala	Leu	Lys	Glu	Ile	Pro	Ser	Gly	Trp	His	Leu	
	1490					1495					1500					
Trp	Arg	Ser	Val	Arg	Ala	Gly	Ile	Met	Pro	Phe	Leu	Lys	Cys	Ser	Ala	
505					1510					1515					1520	
Leu	Phe	Phe	His	Tyr	Leu	Asn	Gly	Val	Pro	Ser	Pro	Pro	Asp	Ile	Gln	
			1525					1530						1535		
Val	Pro	Gly	Thr	Ser	His	Phe	Glu	His	Leu	Cys	Ser	Tyr	Leu	Ser	Leu	
		1540						1545					1550			
Pro	Asn	Asn	Leu	Ile	Cys	Leu	Phe	Gln	Glu	Asn	Ser	Glu	Ile	Met	Asn	
		1555					1560					1565				
Ser	Leu	Ile	Glu	Ser	Trp	Cys	Arg	Asn	Ser	Glu	Val	Lys	Arg	Tyr	Leu	
	1570					1575					1580					
Glu	Gly	Glu	Arg	Asp	Ala	Ile	Arg	Tyr	Pro	Arg	Glu	Ser	Asn	Lys	Leu	
585					1590					1595					1600	
Ile	Asn	Leu	Pro	Glu	Asp	Tyr	Ser	Ser	Leu	Ile	Asn	Gln	Ala	Ser	Asn	
			1605						1610					1615		
Phe	Ser	Cys	Pro	Lys	Ser	Gly	Gly	Asp	Lys	Ser	Arg	Ala	Pro	Thr	Leu	
		1620						1625					1630			
Cys	Leu	Val	Cys	Gly	Ser	Leu	Leu	Cys	Ser	Gln	Ser	Tyr	Cys	Cys	Gln	
		1635					1640					1645				
Thr	Glu	Leu	Glu	Gly	Glu	Asp	Val	Gly	Ala	Cys	Thr	Ala	His	Thr	Tyr	
	1650					1655					1660					
Ser	Cys	Gly	Ser	Gly	Val	Gly	Ile	Phe	Leu	Arg	Val	Arg	Glu	Cys	Gln	
665					1670					1675					1680	

Val Leu Phe Leu Ala Gly Lys Thr Lys Gly Cys Phe Tyr Ser Pro Pro
1685 1690 1695

Tyr Leu Asp Asp Tyr Gly Glu Thr Asp Gln Gly Leu Arg Arg Gly Asn
1700 1705 1710

Pro Leu His Leu Cys Lys Glu Arg Phe Lys Lys Ile Gln Lys Leu Trp
1715 1720 1725

His Gln His Ser Val Thr Glu Glu Ile Gly His Ala Gln Glu Ala Asn
1730 1735 1740

Gln Thr Leu Val Gly Ile Asp Trp Gln His Leu
745 1750 1755

<210> 5
<211> 6089
<212> DNA
<213> Mouse

<220>
<221> CDS
<222> (766) .. (6030)

<400> 5
caagtgtatc atatgccaaag tacgccccct attgacgtca atgacggtaa atggcccgcc 60
tggcattatg cccagtacat gaccttatgg gactttccta cttggcagta catctacgta 120
ttagtcatcg ctattaccat ggtgatgagg ttttggcagt acatcaatgg gcgtaggatag 180
cggtttgact cacgggggatt tccaagtctc caccocattg acgtcaatgg gagtttggtt 240
tggcaccaaa atcaacggga ctttccaaaa tgctgtaaca actccgcccc attgacgcaa 300
atgggacggt ggcgtgtacg gtgggagggtc tatataagca gagctctctg gctaactaga 360
gaaccactg cttactgggt tatcgaaatt aatacgactc actataggga gacccaagct 420
tggtaccgag ctcgatcca ctactcgacc cagcggtccg cggtagtggc tgtccggagt 480
gccaggcctg gggttctcgg tgctcttccc ccgggtcacgg gcgaggaggc gacttgactt 540
ctgggagcgg agcccgggcg cgcgcgcaag cggctgccgt ccccgctgca ggttcgcgtc 600
ccgctttgct cctcgcgcac ctccgctcgg cggcagcccg gacggcccg gactgacggc 660
cccaggacag gggtagcgt cgcggctcgg ggagcagagg cgaagctgag gcccggggag 720
aggcgacagc ggcgagagca cccggggaga ggaggaggag agaag atg gcg tcg gag 777
Met Ala Ser Glu

1

atg gag ccc gag gtg cag gcc atc gac cgc agt ttg ctg gaa tgt tct 825
Met Glu Pro Glu Val Gln Ala Ile Asp Arg Ser Leu Leu Glu Cys Ser
5 10 15 20

gcc gaa gag atc gca ggg aga tgg ctg caa gca acc gac ctg aac aga 873
Ala Glu Glu Ile Ala Gly Arg Trp Leu Gln Ala Thr Asp Leu Asn Arg
25 30 35

gaa gtg tac cag cat tta gcc cac tgt gtg ccc aaa atc tac tgc cgg	921
Glu Val Tyr Gln His Leu Ala His Cys Val Pro Lys Ile Tyr Cys Arg	
40 45 50	
ggc cct aac ccc ttc cct cag aag gaa gac acg ctg gca cag cac atc	969
Gly Pro Asn Pro Phe Pro Gln Lys Glu Asp Thr Leu Ala Gln His Ile	
55 60 65	
ctg ctg gga ccg atg gag tgg tac atc tgc gct gaa gac cct gcg ctg	1017
Leu Leu Gly Pro Met Glu Trp Tyr Ile Cys Ala Glu Asp Pro Ala Leu	
70 75 80	
gga ttt cca aag ctc gag cag gca aac aag cct tct cac ctc tgt ggc	1065
Gly Phe Pro Lys Leu Glu Gln Ala Asn Lys Pro Ser His Leu Cys Gly	
85 90 95 100	
cga gtg ttt aaa gtg ggg gaa cct aca tac tcc tgc aga gac tgt gca	1113
Arg Val Phe Lys Val Gly Glu Pro Thr Tyr Ser Cys Arg Asp Cys Ala	
105 110 115	
gtt gac ccc acc tgt gtt tta tgc atg gag tgc ttc ctg gga agt atc	1161
Val Asp Pro Thr Cys Val Leu Cys Met Glu Cys Phe Leu Gly Ser Ile	
120 125 130	
cat aga gac cat cga tat agg atg acc aca tgc gga gga ggg ggc ttc	1209
His Arg Asp His Arg Tyr Arg Met Thr Thr Ser Gly Gly Gly Gly Phe	
135 140 145	
tgt gac tgt ggt gac act gag gcg tgg aaa gag gga cct tac tgc cag	1257
Cys Asp Cys Gly Asp Thr Glu Ala Trp Lys Glu Gly Pro Tyr Cys Gln	
150 155 160	
aag cac aag ctc agc agc tct gaa gtt gtg gag gag gag gat cct ctt	1305
Lys His Lys Leu Ser Ser Ser Glu Val Val Glu Glu Glu Asp Pro Leu	
165 170 175 180	
gtg cat cta tca gaa gat gtg atc gcc aga act tac aac att ttt gct	1353
Val His Leu Ser Glu Asp Val Ile Ala Arg Thr Tyr Asn Ile Phe Ala	
185 190 195	
att atg ttt cga tat gca gta gat ata ctg acc tgg gaa aaa gaa agt	1401
Ile Met Phe Arg Tyr Ala Val Asp Ile Leu Thr Trp Glu Lys Glu Ser	
200 205 210	
gaa ttg cct gaa gac tta gaa gtg gca gag aag agt gac acc tac tac	1449
Glu Leu Pro Glu Asp Leu Glu Val Ala Glu Lys Ser Asp Thr Tyr Tyr	
215 220 225	
tgc atg ctg ttt aat gat gag gtt cac acc tat gag caa gtc att tat	1497
Cys Met Leu Phe Asn Asp Glu Val His Thr Tyr Glu Gln Val Ile Tyr	
230 235 240	
acc ctt cag aaa gct gtg aac tgt aca cag aag gaa gcc att ggc ttt	1545
Thr Leu Gln Lys Ala Val Asn Cys Thr Gln Lys Glu Ala Ile Gly Phe	
245 250 255 260	
gca act aca gtt gat cga gat ggc cgt agg cct gtc cga tat gga gat	1593
Ala Thr Thr Val Asp Arg Asp Gly Arg Arg Pro Val Arg Tyr Gly Asp	
265 270 275	
ttc cag tac tgt gat caa gca aag aca gtc att gtg agg aac acc agc	1641
Phe Gln Tyr Cys Asp Gln Ala Lys Thr Val Ile Val Arg Asn Thr Ser	
280 285 290	

aga cag acc aag ccg ctc aaa gtt caa gtt atg cac tcc tcc gtg gct	1689
Arg Gln Thr Lys Pro Leu Lys Val Gln Val Met His Ser Ser Val Ala	
295 300 305	
gct cat cag aat ttt ggt ttg aaa gct ctg tcg tgg ctg gga agt gtt	1737
Ala His Gln Asn Phe Gly Leu Lys Ala Leu Ser Trp Leu Gly Ser Val	
310 315 320	
att gga tac tca gat ggc ctt cgc agg att ttg tgt caa gtt gga tta	1785
Ile Gly Tyr Ser Asp Gly Leu Arg Arg Ile Leu Cys Gln Val Gly Leu	
325 330 335 340	
caa gaa ggt cca gat ggc gaa aac tct tct ctg gtc gac aga ctg atg	1833
Gln Glu Gly Pro Asp Gly Glu Asn Ser Ser Leu Val Asp Arg Leu Met	
345 350 355	
ctt aat gat tcc aaa tta tgg aaa ggg gct agg agt gtg tat cac cag	1881
Leu Asn Asp Ser Lys Leu Trp Lys Gly Ala Arg Ser Val Tyr His Gln	
360 365 370	
ttg ttc atg agc agc ctg ctc atg gac ctc aag tat aag aag ctg ttc	1929
Leu Phe Met Ser Ser Leu Leu Met Asp Leu Lys Tyr Lys Lys Leu Phe	
375 380 385	
gcg ctt cga ttt gct aaa aat tac cgg cag ttg cag agg gat ttt atg	1977
Ala Leu Arg Phe Ala Lys Asn Tyr Arg Gln Leu Gln Arg Asp Phe Met	
390 395 400	
gag gat gat cac gag cgg gca gtg tcg gtg act gct ctg tct gtc cag	2025
Glu Asp Asp His Glu Arg Ala Val Ser Val Thr Ala Leu Ser Val Gln	
405 410 415 420	
ttc ttc acc gca ccg acg ctg gcg cga atg ctc ctc aca gaa gag aac	2073
Phe Phe Thr Ala Pro Thr Leu Ala Arg Met Leu Leu Thr Glu Glu Asn	
425 430 435	
ctg atg acc gtt atc att aag gct ttc atg gac cat ttg aaa cac aga	2121
Leu Met Thr Val Ile Ile Lys Ala Phe Met Asp His Leu Lys His Arg	
440 445 450	
gat gcc cag ggc aga ttc cag ttt gaa cgc tac act gcc ctc caa gcc	2169
Asp Ala Gln Gly Arg Phe Gln Phe Glu Arg Tyr Thr Ala Leu Gln Ala	
455 460 465	
ttc aag ttc agg aga gtc cag agc ctc atc tta gat ctc aag tat gta	2217
Phe Lys Phe Arg Arg Val Gln Ser Leu Ile Leu Asp Leu Lys Tyr Val	
470 475 480	
ttg att agc aaa cca acg gag tgg tca gat gag ctg agg cag aag ttc	2265
Leu Ile Ser Lys Pro Thr Glu Trp Ser Asp Glu Leu Arg Gln Lys Phe	
485 490 495 500	
tta caa ggg ttc gat gcc ttc ttg gaa tta ctg aag tgc atg cag gga	2313
Leu Gln Gly Phe Asp Ala Phe Leu Glu Leu Leu Lys Cys Met Gln Gly	
505 510 515	
atg gac ccg atc acg cgt cag gtg gga cag cac att gag atg gag cca	2361
Met Asp Pro Ile Thr Arg Gln Val Gly Gln His Ile Glu Met Glu Pro	
520 525 530	
gag tgg gaa gca gcc ttc aca ctg cag atg aag ctg aca cac gtc atc	2409
Glu Trp Glu Ala Ala Phe Thr Leu Gln Met Lys Leu Thr His Val Ile	
535 540 545	

tca atg gtg cag gac tgg tgt gct ctg gac gaa aaa gtg tta att gaa	2457
Ser Met Val Gln Asp Trp Cys Ala Leu Asp Glu Lys Val Leu Ile Glu	
550 555 560	
gct tac aag aaa tgc ctg gct gtg ctg aca cag tgt cat ggc gga ttt	2505
Ala Tyr Lys Lys Cys Leu Ala Val Leu Thr Gln Cys His Gly Gly Phe	
565 570 575 580	
act gat ggt gaa cag cca atc aca ctc agt att tgt gga cac tcg gtg	2553
Thr Asp Gly Glu Gln Pro Ile Thr Leu Ser Ile Cys Gly His Ser Val	
585 590 595	
gaa acc atc aga tac tgt gtt tcc caa gaa aaa gtt agc att cac ctc	2601
Glu Thr Ile Arg Tyr Cys Val Ser Gln Glu Lys Val Ser Ile His Leu	
600 605 610	
cca att tct cgc ttg ctt gca ggt ttg cat gta ttg tta agc aaa agt	2649
Pro Ile Ser Arg Leu Leu Ala Gly Leu His Val Leu Leu Ser Lys Ser	
615 620 625	
gaa gtg gca tat aaa ttt cca gag ctc cta cct cta agt gaa ctg agc	2697
Glu Val Ala Tyr Lys Phe Pro Glu Leu Leu Pro Leu Ser Glu Leu Ser	
630 635 640	
cca ccc atg ttg ata gaa cat cct ctt aga tgt ctt gtc tta tgt gct	2745
Pro Pro Met Leu Ile Glu His Pro Leu Arg Cys Leu Val Leu Cys Ala	
645 650 655 660	
caa gtg cat gct ggg atg tgg aga aga aat ggc ttc tct cta gta aat	2793
Gln Val His Ala Gly Met Trp Arg Arg Asn Gly Phe Ser Leu Val Asn	
665 670 675	
cag atc tat tac tac cat aat gtg aaa tgc agg cga gag atg ttc gac	2841
Gln Ile Tyr Tyr Tyr His Asn Val Lys Cys Arg Arg Glu Met Phe Asp	
680 685 690	
aag gac ata gtg atg ctt cag aca ggt gtc tcc atg atg gac cca aac	2889
Lys Asp Ile Val Met Leu Gln Thr Gly Val Ser Met Met Asp Pro Asn	
695 700 705	
cac ttc ctg atg atc atg ctc agc cgc ttt gaa ctc tat cag ctc ttc	2937
His Phe Leu Met Ile Met Leu Ser Arg Phe Glu Leu Tyr Gln Leu Phe	
710 715 720	
agc acg cct gac tat ggg aag aga ttc agt tct gag gtt acc cat aag	2985
Ser Thr Pro Asp Tyr Gly Lys Arg Phe Ser Ser Glu Val Thr His Lys	
725 730 735 740	
gac gtc gtt cag cag aac aac act ctg atc gaa gag atg ctc tac ctc	3033
Asp Val Val Gln Gln Asn Asn Thr Leu Ile Glu Glu Met Leu Tyr Leu	
745 750 755	
atc atc atg ctt gtg gga gaa aga ttc aac cct ggg gtt gga cag gtg	3081
Ile Ile Met Leu Val Gly Glu Arg Phe Asn Pro Gly Val Gly Gln Val	
760 765 770	
gct gcc aca gat gaa atc aag agg gag att atc cat cag ttg agc atc	3129
Ala Ala Thr Asp Glu Ile Lys Arg Glu Ile Ile His Gln Leu Ser Ile	
775 780 785	
aag cct atg gct cac agt gag ctg gtg aag tct ctg cct gaa gat gag	3177
Lys Pro Met Ala His Ser Glu Leu Val Lys Ser Leu Pro Glu Asp Glu	
790 795 800	

aac aag gag acc ggc atg gag agc gtc atc gag tcc gtt gca cat ttc	3225
Asn Lys Glu Thr Gly Met Glu Ser Val Ile Glu Ser Val Ala His Phe	
805 810 815 820	
aag aaa cct ggg ctc aca ggg cga ggc atg tat gag ctg aag cca gag	3273
Lys Lys Pro Gly Leu Thr Gly Arg Gly Met Tyr Glu Leu Lys Pro Glu	
825 830 835	
tgt gcc aaa gag ttc aac ctg tat ttt tat cat ttc tcc agg gca gag	3321
Cys Ala Lys Glu Phe Asn Leu Tyr Phe Tyr His Phe Ser Arg Ala Glu	
840 845 850	
cag tcc aag gca gag gaa gct cag cgg aaa ttg aaa aga gaa aat aaa	3369
Gln Ser Lys Ala Glu Glu Ala Gln Arg Lys Leu Lys Arg Glu Asn Lys	
855 860 865	
gaa gat aca gca ctc cct cct ccg gct ttg cca ccg ttc tgc cct ttg	3417
Glu Asp Thr Ala Leu Pro Pro Pro Ala Leu Pro Pro Phe Cys Pro Leu	
870 875 880	
ttc gcg agt ctg gtt aac atc ttg cag tgt gac gtc atg ctg tac atc	3465
Phe Ala Ser Leu Val Asn Ile Leu Gln Cys Asp Val Met Leu Tyr Ile	
885 890 895 900	
atg gga acg atc ctg cag tgg gct gta gag cat cac ggg tct gcc tgg	3513
Met Gly Thr Ile Leu Gln Trp Ala Val Glu His His Gly Ser Ala Trp	
905 910 915	
tca gag tcc atg cta cag agg gtg ctg cat ttg atc ggg atg gct ctc	3561
Ser Glu Ser Met Leu Gln Arg Val Leu His Leu Ile Gly Met Ala Leu	
920 925 930	
cag gaa gag aag cac cac ttg gag aac gcc gtg gaa ggg cac gtg cag	3609
Gln Glu Glu Lys His His Leu Glu Asn Ala Val Glu Gly His Val Gln	
935 940 945	
acc ttc acc ttc aca cag aag att tca aag cct ggt gat gca cca cat	3657
Thr Phe Thr Phe Thr Gln Lys Ile Ser Lys Pro Gly Asp Ala Pro His	
950 955 960	
aac tcc ccg agc atc cta gct atg ctg gag acc ttg cag aac gcc ccc	3705
Asn Ser Pro Ser Ile Leu Ala Met Leu Glu Thr Leu Gln Asn Ala Pro	
965 970 975 980	
tcc ctg gaa gcc cac aag gac atg atc agg tgg ttg cta aag atg ttt	3753
Ser Leu Glu Ala His Lys Asp Met Ile Arg Trp Leu Leu Lys Met Phe	
985 990 995	
aat gca att aag aag ata aga gag tgt tca tcc agc agc cct gtg gcc	3801
Asn Ala Ile Lys Lys Ile Arg Glu Cys Ser Ser Ser Ser Pro Val Ala	
1000 1005 1010	
gag gcg gag gga acc ata atg gag gag agc tca aga gac aag gac aaa	3849
Glu Ala Glu Gly Thr Ile Met Glu Glu Ser Ser Arg Asp Lys Asp Lys	
1015 1020 1025	
gca gag agg aaa aga aaa gcc gag atc gcc aga ctg cgc cgg gag aag	3897
Ala Glu Arg Lys Arg Lys Ala Glu Ile Ala Arg Leu Arg Arg Glu Lys	
1030 1035 1040	
atc atg gcc cag atg tct gag atg cag cgg cac ttc att gac gaa aac	3945
Ile Met Ala Gln Met Ser Glu Met Gln Arg His Phe Ile Asp Glu Asn	
1045 1050 1055 1060	

aaa	gag	ctc	ttc	cag	cag	acc	cta	gag	ctg	gac	acc	tct	gcc	tct	gcc	3993
Lys	Glu	Leu	Phe	Gln	Gln	Thr	Leu	Glu	Leu	Asp	Thr	Ser	Ala	Ser	Ala	
1065				1070				1075								
act	ctt	gac	agc	agc	cct	ccc	gtt	tca	gac	gca	gct	ctt	aca	gca	ctg	4041
Thr	Leu	Asp	Ser	Ser	Pro	Pro	Val	Ser	Asp	Ala	Ala	Leu	Thr	Ala	Leu	
1080				1085				1090								
ggc	cca	gca	cag	aca	cag	gtc	cct	gaa	ccg	aga	cag	ttt	gtc	acc	tgt	4089
Gly	Pro	Ala	Gln	Thr	Gln	Val	Pro	Glu	Pro	Arg	Gln	Phe	Val	Thr	Cys	
1095				1100				1105								
ata	tta	tgt	caa	gag	gag	caa	gag	gtg	act	gtg	gga	agc	agg	gcg	atg	4137
Ile	Leu	Cys	Gln	Glu	Glu	Gln	Glu	Val	Thr	Val	Gly	Ser	Arg	Ala	Met	
1110				1115				1120								
gtc	ttg	gca	gcg	ttt	gtt	cag	agg	tca	acg	gtt	ctg	tca	aaa	gac	agg	4185
Val	Leu	Ala	Ala	Phe	Val	Gln	Arg	Ser	Thr	Val	Leu	Ser	Lys	Asp	Arg	
1125				1130				1135				1140				
acg	aaa	acc	atc	gcg	gac	cca	gaa	aaa	tat	gat	cca	tta	ttc	atg	cac	4233
Thr	Lys	Thr	Ile	Ala	Asp	Pro	Glu	Lys	Tyr	Asp	Pro	Leu	Phe	Met	His	
1145				1150				1155								
ccc	gat	ctg	tct	tgt	ggg	aca	cac	act	ggc	agc	tgt	ggg	cac	gtt	atg	4281
Pro	Asp	Leu	Ser	Cys	Gly	Thr	His	Thr	Gly	Ser	Cys	Gly	His	Val	Met	
1160				1165				1170								
cat	gcc	cat	tgt	tgg	caa	agg	tat	ttt	gat	tcc	gtt	caa	gcc	aag	gag	4329
His	Ala	His	Cys	Trp	Gln	Arg	Tyr	Phe	Asp	Ser	Val	Gln	Ala	Lys	Glu	
1175				1180				1185								
cag	cga	agg	cag	cag	cgg	ctg	cgc	ttg	cac	act	agc	tac	gat	gta	gag	4377
Gln	Arg	Arg	Gln	Gln	Arg	Leu	Arg	Leu	His	Thr	Ser	Tyr	Asp	Val	Glu	
1190				1195				1200								
aat	ggc	gag	ttc	ctc	tgc	ccg	ctc	tgt	gag	tgc	ctg	agc	aac	acg	gtg	4425
Asn	Gly	Glu	Phe	Leu	Cys	Pro	Leu	Cys	Glu	Cys	Leu	Ser	Asn	Thr	Val	
1205				1210				1215				1220				
atc	ccc	ctg	ctg	ctt	cct	ccc	agg	agc	atc	ctc	agc	agg	agg	tta	aat	4473
Ile	Pro	Leu	Leu	Leu	Pro	Pro	Arg	Ser	Ile	Leu	Ser	Arg	Arg	Leu	Asn	
1225				1230				1235								
ttt	tca	gac	caa	cca	gat	ctg	gca	cag	tgg	acg	aga	gca	gta	aca	cag	4521
Phe	Ser	Asp	Gln	Pro	Asp	Leu	Ala	Gln	Trp	Thr	Arg	Ala	Val	Thr	Gln	
1240				1245				1250								
cag	ata	aag	gtg	gtc	cag	atg	ctg	agg	aga	aag	cac	aat	gct	gct	gac	4569
Gln	Ile	Lys	Val	Val	Gln	Met	Leu	Arg	Arg	Lys	His	Asn	Ala	Ala	Asp	
1255				1260				1265								
acg	tct	tct	tca	gag	gac	aca	gaa	gcc	atg	aat	ata	ata	ccg	atc	ccc	4617
Thr	Ser	Ser	Ser	Glu	Asp	Thr	Glu	Ala	Met	Asn	Ile	Ile	Pro	Ile	Pro	
1270				1275				1280								
gaa	ggc	ttc	agg	cct	gat	ttt	tat	cct	agg	aac	cca	tat	tct	gat	agc	4665
Glu	Gly	Phe	Arg	Pro	Asp	Phe	Tyr	Pro	Arg	Asn	Pro	Tyr	Ser	Asp	Ser	
1285				1290				1295				1300				
ata	aaa	gaa														

ctg	aag	gtt	cat	cct	aat	gaa	ggg	gac	ccc	cgt	gtg	ccc	atc	ctg	tgc	4761
Leu	Lys	Val	His	Pro	Asn	Glu	Gly	Asp	Pro	Arg	Val	Pro	Ile	Leu	Cys	
			1320				1325						1330			
tgg	ggg	acc	tgt	gca	tac	acc	atc	cag	agc	ata	gaa	aga	att	ttg	agt	4809
Trp	Gly	Thr	Cys	Ala	Tyr	Thr	Ile	Gln	Ser	Ile	Glu	Arg	Ile	Leu	Ser	
			1335				1340						1345			
gat	gag	gag	aag	cct	ggt	ttt	gga	cct	ctg	cct	tgt	aga	ctg	gac	gac	4857
Asp	Glu	Glu	Lys	Pro	Val	Phe	Gly	Pro	Leu	Pro	Cys	Arg	Leu	Asp	Asp	
			1350				1355						1360			
tgt	ctc	agg	tcg	tta	aca	cgg	ttt	gca	gca	gca	cat	tgg	aca	gtg	gcg	4905
Cys	Leu	Arg	Ser	Leu	Thr	Arg	Phe	Ala	Ala	Ala	His	Trp	Thr	Val	Ala	
			1365				1370						1375			1380
tta	ctt	cct	gtg	gta	caa	gga	cac	ttc	tgt	aaa	ctc	ttt	gca	tcc	ttg	4953
Leu	Leu	Pro	Val	Val	Gln	Gly	His	Phe	Cys	Lys	Leu	Phe	Ala	Ser	Leu	
			1385							1390						1395
gtg	cct	agt	gac	agc	tat	gaa	gac	ctc	cgg	tgc	ata	cta	gac	atc	gac	5001
Val	Pro	Ser	Asp	Ser	Tyr	Glu	Asp	Leu	Pro	Cys	Ile	Leu	Asp	Ile	Asp	
			1400							1405						1410
atg	ttt	cac	ttg	ctg	gtg	ggc	ctg	gtg	ctc	gct	ttc	cca	gct	ctg	cag	5049
Met	Phe	His	Leu	Leu	Val	Gly	Leu	Val	Leu	Ala	Phe	Pro	Ala	Leu	Gln	
			1415				1420						1425			
tgt	cag	gat	ttt	tca	gga	agc	agc	ctg	gcc	act	ggg	gac	ctg	cac	atc	5097
Cys	Gln	Asp	Phe	Ser	Gly	Ser	Ser	Leu	Ala	Thr	Gly	Asp	Leu	His	Ile	
			1430				1435						1440			
ttc	cac	ttg	gtt	acc	atg	gca	cac	atc	gta	cag	atc	tta	ctt	acc	tca	5145
Phe	His	Leu	Val	Thr	Met	Ala	His	Ile	Val	Gln	Ile	Leu	Leu	Thr	Ser	
			1445				1450						1455			1460
tgt	aca	gaa	gag	aat	ggc	atg	gat	caa	gag	aat	ccc	act	ggg	gaa	gaa	5193
Cys	Thr	Glu	Glu	Asn	Gly	Met	Asp	Gln	Glu	Asn	Pro	Thr	Gly	Glu	Glu	
			1465							1470						1475
gaa	ctg	gcc	att	ctc	tct	ttg	cac	aaa	aca	ctt	cac	cag	tat	act	gga	5241
Glu	Leu	Ala	Ile	Leu	Ser	Leu	His	Lys	Thr	Leu	His	Gln	Tyr	Thr	Gly	
			1480							1485						1490
agt	gcc	ttg	aaa	gaa	gcc	ccc	tcc	ggc	tgg	cac	ctg	tgg	agg	agc	gtc	5289
Ser	Ala	Leu	Lys	Glu	Ala	Pro	Ser	Gly	Trp	His	Leu	Trp	Arg	Ser	Val	
			1495				1500						1505			
cgg	gcc	gcc	atc	atg	cct	ttc	ctc	aag	tgc	tct	gct	ttg	ttt	ttc	cac	5337
Arg	Ala	Ala	Ile	Met	Pro	Phe	Leu	Lys	Cys	Ser	Ala	Leu	Phe	Phe	His	
			1510				1515						1520			
tat	tta	aat	gga	gtc	cgg	gcc	cct	cca	gac	ctt	caa	gtt	tct	gga	aca	5385
Tyr	Leu	Asn	Gly	Val	Pro	Ala	Pro	Pro	Asp	Leu	Gln	Val	Ser	Gly	Thr	
			1525				1530						1535			1540
agc	cat	ttt	gaa	cac	tta	tgt</										

agt tgg tgc cag aac agt gaa gtt aaa cgg tat cta aat ggc gag aga 5529
 Ser Trp Cys Gln Asn Ser Glu Val Lys Arg Tyr Leu Asn Gly Glu Arg
 1575 1580 1585

 gga gcg ata agc tac ccc aga gga gct aac aaa ctg ata gac ctt cca 5577
 Gly Ala Ile Ser Tyr Pro Arg Gly Ala Asn Lys Leu Ile Asp Leu Pro
 1590 1595 1600

 gag gat tac agc agc ctc att aac caa gca tcc aat ttc tcg tgc ccc 5625
 Glu Asp Tyr Ser Ser Leu Ile Asn Gln Ala Ser Asn Phe Ser Cys Pro
 1605 1610 1615 1620

 aaa tca ggt ggc gac aag agc aga gct cct act ctg tgc ctc gtg tgt 5673
 Lys Ser Gly Gly Asp Lys Ser Arg Ala Pro Thr Leu Cys Leu Val Cys
 1625 1630 1635

 ggg agt ctc ctc tgc tct cag agt tac tgc tgc caa gct gag ctg gag 5721
 Gly Ser Leu Leu Cys Ser Gln Ser Tyr Cys Cys Gln Ala Glu Leu Glu
 1640 1645 1650

 ggt gag gac gtc gga gcc tgc aca gca cac acc tac tcc tgc ggc tcc 5769
 Gly Glu Asp Val Gly Ala Cys Thr Ala His Thr Tyr Ser Cys Gly Ser
 1655 1660 1665

 ggg gcc ggc atc ttc ctg aga gtg cgg gaa tgt cag gtg cta ttt tta 5817
 Gly Ala Gly Ile Phe Leu Arg Val Arg Glu Cys Gln Val Leu Phe Leu
 1670 1675 1680

 gct ggc aaa acc aaa gga tgt ttt tat tct cct cct tac ctt gac gac 5865
 Ala Gly Lys Thr Lys Gly Cys Phe Tyr Ser Pro Pro Tyr Leu Asp Asp
 1685 1690 1695 1700

 tat gga gag acc gac cag gga ctc aga cga gga aat cct tta cat tta 5913
 Tyr Gly Glu Thr Asp Gln Gly Leu Arg Arg Gly Asn Pro Leu His Leu
 1705 1710 1715

 tgc caa gag cgg ttt cga aag atc cag aag ctc tgg cag cag cat agt 5961
 Cys Gln Glu Arg Phe Arg Lys Ile Gln Lys Leu Trp Gln Gln His Ser
 1720 1725 1730

 atc aca gag gag atc gga cac gcg cag gag gct aac cag acc ctg gtc 6009
 Ile Thr Glu Glu Ile Gly His Ala Gln Glu Ala Asn Gln Thr Leu Val
 1735 1740 1745

 gga att gac tgg cag cat tta taatcgctcc tctactaaaa acttgacttg 6060
 Gly Ile Asp Trp Gln His Leu
 1750 1755

 gagttttgta acacagctgg cttttccag 6089

<210> 6
 <211> 1755
 <212> PRT
 <213> Mouse

<400> 6
 Met Ala Ser Glu Met Glu Pro Glu Val Gln Ala Ile Asp Arg Ser Leu
 1 5 10 15
 Leu Glu Cys Ser Ala Glu Glu Ile Ala Gly Arg Trp Leu Gln Ala Thr
 20 25 30

Asp	Leu	Asn	Arg	Glu	Val	Tyr	Gln	His	Leu	Ala	His	Cys	Val	Pro	Lys
		35					40					45			
Ile	Tyr	Cys	Arg	Gly	Pro	Asn	Pro	Phe	Pro	Gln	Lys	Glu	Asp	Thr	Leu
	50					55					60				
Ala	Gln	His	Ile	Leu	Leu	Gly	Pro	Met	Glu	Trp	Tyr	Ile	Cys	Ala	Glu
65					70					75					80
Asp	Pro	Ala	Leu	Gly	Phe	Pro	Lys	Leu	Glu	Gln	Ala	Asn	Lys	Pro	Ser
				85					90					95	
His	Leu	Cys	Gly	Arg	Val	Phe	Lys	Val	Gly	Glu	Pro	Thr	Tyr	Ser	Cys
			100				105						110		
Arg	Asp	Cys	Ala	Val	Asp	Pro	Thr	Cys	Val	Leu	Cys	Met	Glu	Cys	Phe
		115					120					125			
Leu	Gly	Ser	Ile	His	Arg	Asp	His	Arg	Tyr	Arg	Met	Thr	Thr	Ser	Gly
	130					135					140				
Gly	Gly	Gly	Phe	Cys	Asp	Cys	Gly	Asp	Thr	Glu	Ala	Trp	Lys	Glu	Gly
145					150					155					160
Pro	Tyr	Cys	Gln	Lys	His	Lys	Leu	Ser	Ser	Ser	Glu	Val	Val	Glu	Glu
				165					170					175	
Glu	Asp	Pro	Leu	Val	His	Leu	Ser	Glu	Asp	Val	Ile	Ala	Arg	Thr	Tyr
			180					185					190		
Asn	Ile	Phe	Ala	Ile	Met	Phe	Arg	Tyr	Ala	Val	Asp	Ile	Leu	Thr	Trp
		195					200					205			
Glu	Lys	Glu	Ser	Glu	Leu	Pro	Glu	Asp	Leu	Glu	Val	Ala	Glu	Lys	Ser
	210					215					220				
Asp	Thr	Tyr	Tyr	Cys	Met	Leu	Phe	Asn	Asp	Glu	Val	His	Thr	Tyr	Glu
225					230					235					240
Gln	Val	Ile	Tyr	Thr	Leu	Gln	Lys	Ala	Val	Asn	Cys	Thr	Gln	Lys	Glu
				245					250					255	
Ala	Ile	Gly	Phe	Ala	Thr	Thr	Val	Asp	Arg	Asp	Gly	Arg	Arg	Pro	Val
			260					265					270		
Arg	Tyr	Gly	Asp	Phe	Gln	Tyr	Cys	Asp	Gln	Ala	Lys	Thr	Val	Ile	Val
		275					280					285			
Arg	Asn	Thr	Ser	Arg	Gln	Thr	Lys	Pro	Leu	Lys	Val	Gln	Val	Met	His
	290					295					300				
Ser	Ser	Val	Ala	Ala	His	Gln	Asn	Phe	Gly	Leu	Lys	Ala	Leu	Ser	Trp
305					310					315					320
Leu	Gly	Ser	Val	Ile	Gly	Tyr	Ser	Asp	Gly	Leu	Arg	Arg	Ile	Leu	Cys
				325					330					335	
Gln	Val	Gly	Leu	Gln	Glu	Gly	Pro	Asp	Gly	Glu	Asn	Ser	Ser	Leu	Val
			340					345					350		
Asp	Arg	Leu	Met	Leu	Asn	Asp	Ser	Lys	Leu	Trp	Lys	Gly	Ala	Arg	Ser
		355					360					365			

Val 370	Tyr	His	Gln	Leu	Phe	Met 375	Ser	Ser	Leu	Leu	Met 380	Asp	Leu	Lys	Tyr
Lys 385	Lys	Leu	Phe	Ala	Leu 390	Arg	Phe	Ala	Lys	Asn 395	Tyr	Arg	Gln	Leu	Gln 400
Arg	Asp	Phe	Met	Glu 405	Asp	Asp	His	Glu	Arg 410	Ala	Val	Ser	Val	Thr 415	Ala
Leu	Ser	Val	Gln 420	Phe	Phe	Thr	Ala	Pro 425	Thr	Leu	Ala	Arg	Met 430	Leu	Leu
Thr	Glu	Glu 435	Asn	Leu	Met	Thr	Val 440	Ile	Ile	Lys	Ala	Phe 445	Met	Asp	His
Leu	Lys 450	His	Arg	Asp	Ala	Gln 455	Gly	Arg	Phe	Gln	Phe 460	Glu	Arg	Tyr	Thr
Ala 465	Leu	Gln	Ala	Phe	Lys 470	Phe	Arg	Arg	Val	Gln 475	Ser	Leu	Ile	Leu	Asp 480
Leu	Lys	Tyr	Val	Leu 485	Ile	Ser	Lys	Pro	Thr 490	Glu	Trp	Ser	Asp	Glu 495	Leu
Arg	Gln	Lys	Phe 500	Leu	Gln	Gly	Phe	Asp 505	Ala	Phe	Leu	Glu	Leu 510	Leu	Lys
Cys	Met	Gln 515	Gly	Met	Asp	Pro	Ile 520	Thr	Arg	Gln	Val	Gly 525	Gln	His	Ile
Glu 530	Met	Glu	Pro	Glu	Trp	Glu 535	Ala	Ala	Phe	Thr	Leu 540	Gln	Met	Lys	Leu
Thr 545	His	Val	Ile	Ser	Met 550	Val	Gln	Asp	Trp	Cys 555	Ala	Leu	Asp	Glu	Lys 560
Val	Leu	Ile	Glu	Ala 565	Tyr	Lys	Lys	Cys	Leu 570	Ala	Val	Leu	Thr	Gln 575	Cys
His	Gly	Gly	Phe 580	Thr	Asp	Gly	Glu	Gln 585	Pro	Ile	Thr	Leu	Ser 590	Ile	Cys
Gly	His	Ser 595	Val	Glu	Thr	Ile	Arg 600	Tyr	Cys	Val	Ser	Gln 605	Glu	Lys	Val
Ser 610	Ile	His	Leu	Pro	Ile	Ser 615	Arg	Leu	Leu	Ala	Gly 620	Leu	His	Val	Leu
Leu 625	Ser	Lys	Ser	Glu	Val 630	Ala	Tyr	Lys	Phe	Pro 635	Glu	Leu	Leu	Pro	Leu 640
Ser	Glu	Leu	Ser	Pro 645	Pro	Met	Leu	Ile	Glu 650	His	Pro	Leu	Arg	Cys 655	Leu
Val	Leu	Cys	Ala 660	Gln	Val	His	Ala	Gly 665	Met	Trp	Arg	Arg	Asn 670	Gly	Phe
Ser	Leu	Val	Asn 675	Gln	Ile	Tyr	Tyr 680	Tyr	His	Asn	Val	Lys 685	Cys	Arg	Arg
Glu	Met 690	Phe	Asp	Lys	Asp	Ile 695	Val	Met	Leu	Gln	Thr 700	Gly	Val	Ser	Met

Met	Asp	Pro	Asn	His	Phe	Leu	Met	Ile	Met	Leu	Ser	Arg	Phe	Glu	Leu
705					710					715					720
Tyr	Gln	Leu	Phe	Ser	Thr	Pro	Asp	Tyr	Gly	Lys	Arg	Phe	Ser	Ser	Glu
				725					730					735	
Val	Thr	His	Lys	Asp	Val	Val	Gln	Gln	Asn	Asn	Thr	Leu	Ile	Glu	Glu
			740					745					750		
Met	Leu	Tyr	Leu	Ile	Ile	Met	Leu	Val	Gly	Glu	Arg	Phe	Asn	Pro	Gly
		755					760					765			
Val	Gly	Gln	Val	Ala	Ala	Thr	Asp	Glu	Ile	Lys	Arg	Glu	Ile	Ile	His
	770					775					780				
Gln	Leu	Ser	Ile	Lys	Pro	Met	Ala	His	Ser	Glu	Leu	Val	Lys	Ser	Leu
785					790					795					800
Pro	Glu	Asp	Glu	Asn	Lys	Glu	Thr	Gly	Met	Glu	Ser	Val	Ile	Glu	Ser
				805					810					815	
Val	Ala	His	Phe	Lys	Lys	Pro	Gly	Leu	Thr	Gly	Arg	Gly	Met	Tyr	Glu
			820					825					830		
Leu	Lys	Pro	Glu	Cys	Ala	Lys	Glu	Phe	Asn	Leu	Tyr	Phe	Tyr	His	Phe
		835					840					845			
Ser	Arg	Ala	Glu	Gln	Ser	Lys	Ala	Glu	Glu	Ala	Gln	Arg	Lys	Leu	Lys
	850					855					860				
Arg	Glu	Asn	Lys	Glu	Asp	Thr	Ala	Leu	Pro	Pro	Pro	Ala	Leu	Pro	Pro
865					870					875					880
Phe	Cys	Pro	Leu	Phe	Ala	Ser	Leu	Val	Asn	Ile	Leu	Gln	Cys	Asp	Val
				885					890					895	
Met	Leu	Tyr	Ile	Met	Gly	Thr	Ile	Leu	Gln	Trp	Ala	Val	Glu	His	His
			900					905					910		
Gly	Ser	Ala	Trp	Ser	Glu	Ser	Met	Leu	Gln	Arg	Val	Leu	His	Leu	Ile
		915					920					925			
Gly	Met	Ala	Leu	Gln	Glu	Glu	Lys	His	His	Leu	Glu	Asn	Ala	Val	Glu
	930					935					940				
Gly	His	Val	Gln	Thr	Phe	Thr	Phe	Thr	Gln	Lys	Ile	Ser	Lys	Pro	Gly
945					950					955					960
Asp	Ala	Pro	His	Asn	Ser	Pro	Ser	Ile	Leu	Ala	Met	Leu	Glu	Thr	Leu
				965					970					975	
Gln	Asn	Ala	Pro	Ser	Leu	Glu	Ala	His	Lys	Asp	Met	Ile	Arg	Trp	Leu
			980					985					990		
Leu	Lys	Met	Phe	Asn	Ala	Ile	Lys	Lys	Ile	Arg	Glu	Cys	Ser	Ser	Ser
		995					1000					1005			
Ser	Pro	Val	Ala	Glu	Ala	Glu	Gly	Thr	Ile	Met	Glu	Glu	Ser	Ser	Arg
	1010					1015					1020				
Asp	Lys	Asp	Lys	Ala	Glu	Arg	Lys	Arg	Lys	Ala	Glu	Ile	Ala	Arg	Leu
025					1030					1035					1040

Arg Arg Glu Lys Ile Met Ala Gln Met Ser Glu Met Gln Arg His Phe
1045 1050 1055

Ile Asp Glu Asn Lys Glu Leu Phe Gln Gln Thr Leu Glu Leu Asp Thr
1060 1065 1070

Ser Ala Ser Ala Thr Leu Asp Ser Ser Pro Pro Val Ser Asp Ala Ala
1075 1080 1085

Leu Thr Ala Leu Gly Pro Ala Gln Thr Gln Val Pro Glu Pro Arg Gln
1090 1095 1100

Phe Val Thr Cys Ile Leu Cys Gln Glu Glu Gln Glu Val Thr Val Gly
1105 1110 1115 1120

Ser Arg Ala Met Val Leu Ala Ala Phe Val Gln Arg Ser Thr Val Leu
1125 1130 1135

Ser Lys Asp Arg Thr Lys Thr Ile Ala Asp Pro Glu Lys Tyr Asp Pro
1140 1145 1150

Leu Phe Met His Pro Asp Leu Ser Cys Gly Thr His Thr Gly Ser Cys
1155 1160 1165

Gly His Val Met His Ala His Cys Trp Gln Arg Tyr Phe Asp Ser Val
1170 1175 1180

Gln Ala Lys Glu Gln Arg Arg Gln Gln Arg Leu Arg Leu His Thr Ser
1185 1190 1195 1200

Tyr Asp Val Glu Asn Gly Glu Phe Leu Cys Pro Leu Cys Glu Cys Leu
1205 1210 1215

Ser Asn Thr Val Ile Pro Leu Leu Leu Pro Pro Arg Ser Ile Leu Ser
1220 1225 1230

Arg Arg Leu Asn Phe Ser Asp Gln Pro Asp Leu Ala Gln Trp Thr Arg
1235 1240 1245

Ala Val Thr Gln Gln Ile Lys Val Val Gln Met Leu Arg Arg Lys His
1250 1255 1260

Asn Ala Ala Asp Thr Ser Ser Ser Glu Asp Thr Glu Ala Met Asn Ile
1265 1270 1275 1280

Ile Pro Ile Pro Glu Gly Phe Arg Pro Asp Phe Tyr Pro Arg Asn Pro
1285 1290 1295

Tyr Ser Asp Ser Ile Lys Glu Met Leu Thr Thr Phe Gly Thr Ala Ala
1300 1305 1310

Tyr Lys Val Gly Leu Lys Val His Pro Asn Glu Gly Asp Pro Arg Val
1315 1320 1325

Pro Ile Leu Cys Trp Gly Thr Cys Ala Tyr Thr Ile Gln Ser Ile Glu
1330 1335 1340

Arg Ile Leu Ser Asp Glu Glu Lys Pro Val Phe Gly Pro Leu Pro Cys
1345 1350 1355 1360

Arg Leu Asp Asp Cys Leu Arg Ser Leu Thr Arg Phe Ala Ala Ala His
1365 1370 1375

Trp Thr Val Ala Leu Leu Pro Val Val Gln Gly His Phe Cys Lys Leu
1380 1385 1390

Phe Ala Ser Leu Val Pro Ser Asp Ser Tyr Glu Asp Leu Pro Cys Ile
1395 1400 1405

Leu Asp Ile Asp Met Phe His Leu Leu Val Gly Leu Val Leu Ala Phe
1410 1415 1420

Pro Ala Leu Gln Cys Gln Asp Phe Ser Gly Ser Ser Leu Ala Thr Gly
425 1430 1435 1440

Asp Leu His Ile Phe His Leu Val Thr Met Ala His Ile Val Gln Ile
1445 1450 1455

Leu Leu Thr Ser Cys Thr Glu Glu Asn Gly Met Asp Gln Glu Asn Pro
1460 1465 1470

Thr Gly Glu Glu Glu Leu Ala Ile Leu Ser Leu His Lys Thr Leu His
1475 1480 1485

Gln Tyr Thr Gly Ser Ala Leu Lys Glu Ala Pro Ser Gly Trp His Leu
1490 1495 1500

Trp Arg Ser Val Arg Ala Ala Ile Met Pro Phe Leu Lys Cys Ser Ala
505 1510 1515 1520

Leu Phe Phe His Tyr Leu Asn Gly Val Pro Ala Pro Pro Asp Leu Gln
1525 1530 1535

Val Ser Gly Thr Ser His Phe Glu His Leu Cys Asn Tyr Leu Ser Leu
1540 1545 1550

Pro Thr Asn Leu Ile His Leu Phe Gln Glu Asn Ser Asp Ile Met Asn
1555 1560 1565

Ser Leu Ile Glu Ser Trp Cys Gln Asn Ser Glu Val Lys Arg Tyr Leu
1570 1575 1580

Asn Gly Glu Arg Gly Ala Ile Ser Tyr Pro Arg Gly Ala Asn Lys Leu
585 1590 1595 1600

Ile Asp Leu Pro Glu Asp Tyr Ser Ser Leu Ile Asn Gln Ala Ser Asn
1605 1610 1615

Phe Ser Cys Pro Lys Ser Gly Gly Asp Lys Ser Arg Ala Pro Thr Leu
1620 1625 1630

Cys Leu Val Cys Gly Ser Leu Leu Cys Ser Gln Ser Tyr Cys Cys Gln
1635 1640 1645

Ala Glu Leu Glu Gly Glu Asp Val Gly Ala Cys Thr Ala His Thr Tyr
1650 1655 1660

Ser Cys Gly Ser Gly Ala Gly Ile Phe Leu Arg Val Arg Glu Cys Gln
665 1670 1675 1680

Val Leu Phe Leu Ala Gly Lys Thr Lys Gly Cys Phe Tyr Ser Pro Pro
1685 1690 1695

Tyr Leu Asp Asp Tyr Gly Glu Thr Asp Gln Gly Leu Arg Arg Gly Asn
1700 1705 1710

Pro Leu His Leu Cys Gln Glu Arg Phe Arg Lys Ile Gln Lys Leu Trp
1715 1720 1725

Gln Gln His Ser Ile Thr Glu Glu Ile Gly His Ala Gln Glu Ala Asn
1730 1735 1740

Gln Thr Leu Val Gly Ile Asp Trp Gln His Leu
1745 1750 1755

<210> 7
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer
-2282-91

<400> 7
ctgctcgagt ctgcgtcaaa c 21

<210> 8
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer -
2385-35

<400> 8
tctcgatatg ttgcagcctt gcta 24

<210> 9
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer 2282-94

<400> 9
gtatgaactt gccgaggctt tta 23

<210> 10
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer 2294-37

<400> 10
caataactttc ccagccctca gaa 23

<210> 11
<211> 18
<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer 2380-88

<400> 11

atggcgtcgc tagagcca

18

<210> 12

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer 2378-32

<400> 12

caaagcggct gagcatgata atc

23

<210> 13

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer 2381-48

<400> 13

tgaacagcca atcacactaa gca

23

<210> 14

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer 2385-94

<400> 14

ttataaatgc caaatgccaa

20

<210> 15

<211> 1757

<212> PRT

<213> Mouse

<400> 15

Met Ala Asp Glu Glu Met Asp Gly Ala Glu Arg Met Asp Val Ser Pro
1 5 10 15

Glu Pro Pro Leu Ala Pro Gln Arg Pro Ala Ser Trp Trp Asp Gln Gln
20 25 30

Val Asp Phe Tyr Thr Ala Phe Leu His His Leu Ala Gln Leu Val Pro
35 40 45

Glu Ile Tyr Phe Ala Glu Met Asp Pro Asp Leu Glu Lys Gln Glu Glu
50 55 60

Ser 65	Val	Gln	Met	Ser	Ile 70	Leu	Thr	Pro	Leu	Glu 75	Trp	Tyr	Leu	Phe	Gly 80
Glu	Asp	Pro	Asp	Ile 85	Cys	Leu	Glu	Lys	Leu 90	Lys	His	Ser	Gly	Ala 95	Phe
Gln	Leu	Cys	Gly 100	Lys	Val	Phe	Lys	Ser 105	Gly	Glu	Thr	Thr	Tyr 110	Ser	Cys
Arg	Asp	Cys 115	Ala	Ile	Asp	Pro	Thr 120	Cys	Val	Leu	Cys	Met 125	Asp	Cys	Phe
Gln	Ser 130	Ser	Val	His	Lys 135	Asn	His	Arg	Tyr	Lys 140	Met	His	Thr	Ser	Thr
Gly 145	Gly	Gly	Phe	Cys	Asp 150	Cys	Gly	Asp	Thr	Glu 155	Ala	Trp	Lys	Thr	Gly 160
Pro	Phe	Cys	Val	Asp 165	His	Glu	Pro	Gly 170	Arg	Ala	Gly	Thr	Thr	Lys 175	Glu
Ser	Leu	His	Cys 180	Pro	Leu	Asn	Glu	Glu 185	Val	Ile	Ala	Gln	Ala	Arg	Arg
Ile	Phe 195	Pro	Ser	Val	Ile	Lys	Tyr 200	Ile	Val	Glu	Met 205	Thr	Ile	Trp	Glu
Glu 210	Glu	Lys	Glu	Leu	Pro	Pro 215	Glu	Leu	Gln	Ile	Arg 220	Glu	Lys	Asn	Glu
Arg 225	Tyr	Tyr	Cys	Val	Leu 230	Phe	Asn	Asp	Glu	His 235	His	Ser	Tyr	Asp	His 240
Val	Ile	Tyr	Ser	Leu 245	Gln	Arg	Ala	Leu	Asp 250	Cys	Glu	Leu	Ala	Glu 255	Ala
Gln	Leu	His	Thr 260	Thr	Ala	Ile	Asp	Lys 265	Glu	Gly	Arg	Arg	Ala 270	Val	Lys
Ala	Gly 275	Val	Tyr	Ala	Thr	Cys	Gln 280	Glu	Ala	Lys	Glu	Asp 285	Ile	Lys	Ser
His 290	Ser	Glu	Asn	Val	Ser	Gln 295	His	Pro	Leu	His 300	Val	Glu	Val	Leu	His
Ser 305	Val	Val	Met	Ala	His 310	Gln	Lys	Phe	Ala	Leu 315	Arg	Leu	Gly	Ser	Trp 320
Met	Asn	Lys	Ile	Met 325	Ser	Tyr	Ser	Ser	Asp 330	Phe	Arg	Gln	Ile	Phe 335	Cys
Gln	Ala	Cys	Leu 340	Val	Glu	Glu	Pro	Gly 345	Ser	Glu	Asn	Pro	Cys 350	Leu	Ile
Ser	Arg 355	Leu	Met	Leu	Trp	Asp	Ala 360	Lys	Leu	Tyr	Lys	Gly 365	Ala	Arg	Lys
Ile 370	Leu	His	Glu	Leu	Ile	Phe 375	Ser	Ser	Phe	Phe 380	Met	Glu	Met	Glu	Tyr
Lys 385	Lys	Leu	Phe	Ala	Met 390	Glu	Phe	Val	Lys	Tyr 395	Tyr	Lys	Gln	Leu	Gln 400

Lys	Glu	Tyr	Ile	Ser	Asp	Asp	His	Glu	Arg	Ser	Ile	Ser	Ile	Thr	Ala	405	410	415
Leu	Ser	Val	Gln	Met	Leu	Thr	Val	Pro	Thr	Leu	Ala	Arg	His	Leu	Ile	420	425	430
Glu	Glu	Gln	Asn	Val	Ile	Ser	Val	Ile	Thr	Glu	Thr	Leu	Leu	Glu	Val	435	440	445
Leu	Pro	Glu	Tyr	Leu	Asp	Arg	Asn	Asn	Lys	Phe	Asn	Phe	Gln	Gly	Tyr	450	455	460
Ser	Gln	Asp	Lys	Leu	Gly	Arg	Val	Tyr	Ala	Val	Ile	Cys	Asp	Leu	Lys	465	470	475
Tyr	Ile	Leu	Ile	Ser	Lys	Pro	Val	Ile	Trp	Thr	Glu	Arg	Leu	Arg	Ala	485	490	495
Gln	Phe	Leu	Glu	Gly	Phe	Arg	Ser	Phe	Leu	Lys	Ile	Leu	Thr	Cys	Met	500	505	510
Gln	Gly	Met	Glu	Glu	Ile	Arg	Arg	Gln	Val	Gly	Gln	His	Ile	Glu	Val	515	520	525
Asp	Pro	Asp	Trp	Glu	Ala	Ala	Ile	Ala	Ile	Gln	Met	Gln	Leu	Lys	Asn	530	535	540
Ile	Leu	Leu	Met	Phe	Gln	Glu	Trp	Cys	Ala	Cys	Asp	Glu	Asp	Leu	Leu	545	550	555
Leu	Val	Ala	Tyr	Lys	Glu	Cys	His	Lys	Ala	Val	Met	Arg	Cys	Ser	Thr	565	570	575
Asn	Phe	Met	Ser	Ser	Thr	Lys	Thr	Val	Val	Gln	Leu	Cys	Gly	His	Ser	580	585	590
Leu	Glu	Thr	Lys	Ser	Tyr	Lys	Val	Ser	Glu	Asp	Leu	Val	Ser	Ile	His	595	600	605
Leu	Pro	Leu	Ser	Arg	Thr	Leu	Ala	Gly	Leu	His	Val	Arg	Leu	Ser	Arg	610	615	620
Leu	Gly	Ala	Ile	Ser	Arg	Leu	His	Glu	Phe	Val	Pro	Phe	Asp	Ser	Phe	625	630	635
Gln	Val	Glu	Val	Leu	Val	Glu	Tyr	Pro	Leu	Arg	Cys	Leu	Val	Leu	Val	645	650	655
Ala	Gln	Val	Val	Ala	Glu	Met	Trp	Arg	Arg	Asn	Gly	Leu	Ser	Leu	Ile	660	665	670
Ser	Gln	Val	Phe	Tyr	Tyr	Gln	Asp	Val	Lys	Cys	Arg	Glu	Glu	Met	Tyr	675	680	685
Asp	Lys	Asp	Ile	Ile	Met	Leu	Gln	Ile	Gly	Ala	Ser	Ile	Met	Asp	Pro	690	695	700
Asn	Lys	Phe	Leu	Leu	Leu	Val	Leu	Gln	Arg	Tyr	Glu	Leu	Thr	Asp	Ala	705	710	715
Phe	Asn	Lys	Thr	Ile	Ser	Thr	Lys	Asp	Gln	Asp	Leu	Ile	Lys	Gln	Tyr	725	730	735

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

Asn	Thr	Leu	Ile	Glu	Glu	Met	Leu	Gln	Val	Leu	Ile	Tyr	Ile	Val	Gly
			740				745						750		
Glu	Arg	Tyr	Val	Pro	Gly	Val	Gly	Asn	Val	Thr	Arg	Glu	Glu	Val	Ile
			755				760						765		
Met	Arg	Glu	Ile	Thr	His	Leu	Leu	Cys	Ile	Glu	Pro	Met	Pro	His	Ser
			770				775						780		
Ala	Ile	Ala	Arg	Asn	Leu	Pro	Glu	Asn	Glu	Asn	Asn	Glu	Thr	Gly	Leu
			785							795			800		
Glu	Asn	Val	Ile	Asn	Lys	Val	Ala	Thr	Phe	Lys	Lys	Pro	Gly	Val	Ser
			805						810			815			
Gly	His	Gly	Val	Tyr	Glu	Leu	Lys	Asp	Glu	Ser	Leu	Lys	Asp	Phe	Asn
			820						825			830			
Met	Tyr	Phe	Tyr	His	Tyr	Ser	Lys	Thr	Gln	His	Ser	Lys	Ala	Glu	His
			835			840						845			
Met	Gln	Lys	Lys	Arg	Arg	Lys	Gln	Glu	Asn	Lys	Asp	Glu	Ala	Leu	Pro
			850			855						860			
Pro	Pro	Pro	Pro	Pro	Glu	Phe	Cys	Pro	Ala	Phe	Ser	Lys	Val	Val	Asn
			865			870			875			880			
Leu	Leu	Ser	Cys	Asp	Val	Met	Ile	Tyr	Ile	Leu	Arg	Thr	Ile	Phe	Glu
			885						890			895			
Arg	Ala	Val	Asp	Thr	Glu	Ser	Asn	Leu	Trp	Thr	Glu	Gly	Met	Leu	Gln
			900			905						910			
Met	Ala	Phe	His	Ile	Leu	Ala	Leu	Gly	Leu	Leu	Glu	Glu	Lys	Gln	Gln
			915			920						925			
Leu	Gln	Lys	Ala	Pro	Glu	Glu	Glu	Val	Ala	Phe	Asp	Phe	Tyr	His	Lys
			930			935			940						
Ala	Ser	Arg	Leu	Gly	Ser	Ser	Ala	Met	Asn	Ala	Gln	Asn	Ile	Gln	Met
			945						955			960			
Leu	Leu	Glu	Arg	Leu	Lys	Gly	Ile	Pro	Gln	Leu	Glu	Gly	Gln	Lys	Asp
			965						970			975			
Met	Ile	Thr	Trp	Ile	Leu	Gln	Met	Phe	Asp	Thr	Val	Lys	Arg	Leu	Arg
			980			985						990			
Glu	Lys	Ser	Cys	Leu	Val	Val	Ala	Thr	Thr	Ser	Gly	Leu	Glu	Cys	Ile
			995			1000						1005			
Lys	Ser	Glu	Glu	Ile	Thr	His	Asp	Lys	Glu	Lys	Ala	Glu	Arg	Lys	Arg
			1010			1015						1020			
Lys	Ala	Glu	Ala	Ala	Arg	Leu	His	Arg	Gln	Lys	Ile	Met	Ala	Gln	Met
			1025			1030			1035			1040			
Ser	Ala	Leu	Gln	Lys	Asn	Phe	Ile	Glu	Thr	His	Lys	Leu	Met	Tyr	Asp
			1045			1050						1055			
Asn	Thr	Ser	Glu	Val	Thr	Gly	Lys	Glu	Asp	Ser	Ile	Met	Glu	Glu	Glu
			1060			1065						1070			

Ser Thr Ser Ala Val Ser Glu Ala Ser Arg Ile Ala Leu Gly Pro Lys
1075 1080 1085

Arg Gly Pro Ala Val Thr Glu Lys Glu Val Leu Thr Cys Ile Leu Cys
1090 1095 1100

Gln Glu Glu Gln Glu Val Lys Leu Glu Asn Asn Ala Met Val Leu Ser
1105 1110 1115 1120

Ala Cys Val Gln Lys Ser Thr Ala Leu Thr Gln His Arg Gly Lys Pro
1125 1130 1135

Val Asp His Leu Gly Glu Thr Leu Asp Pro Leu Phe Met Asp Pro Asp
1140 1145 - 1150

Leu Ala His Gly Thr Tyr Thr Gly Ser Cys Gly His Val Met His Ala
1155 1160 1165

Val Cys Trp Gln Lys Tyr Phe Glu Ala Val Gln Leu Ser Ser Gln Gln
1170 1175 1180

Arg Ile His Val Asp Leu Phe Asp Leu Glu Ser Gly Glu Tyr Leu Cys
1185 1190 1195 1200

Pro Leu Cys Lys Ser Leu Cys Asn Thr Val Ile Pro Ile Ile Pro Leu
1205 1210 1215

Gln Pro Gln Lys Ile Asn Ser Glu Asn Ala Glu Ala Leu Ala Gln Leu
1220 1225 1230

Leu Thr Leu Ala Arg Trp Ile Gln Thr Val Leu Ala Arg Ile Ser Gly
1235 1240 1245

Tyr Asn Ile Lys His Ala Lys Gly Glu Ala Pro Ala Val Pro Val Leu
1250 1255 1260

Phe Asn Gln Gly Met Gly Asp Ser Thr Phe Glu Phe His Ser Ile Leu
1265 1270 1275 1280

Ser Phe Gly Val Gln Ser Ser Val Lys Tyr Ser Asn Ser Ile Lys Glu
1285 1290 1295

Met Val Ile Leu Phe Ala Thr Thr Ile Tyr Arg Ile Gly Leu Lys Val
1300 1305 1310

Pro Pro Asp Glu Leu Asp Pro Arg Val Pro Met Met Thr Trp Ser Thr
1315 1320 1325

Cys Ala Phe Thr Ile Gln Ala Ile Glu Asn Leu Leu Gly Asp Glu Gly
1330 1335 1340

Lys Pro Leu Phe Gly Ala Leu Gln Asn Arg Gln His Ser Gly Leu Lys
1345 1350 1355 1360

Ala Leu Met Gln Phe Ala Val Ala Gln Arg Ala Thr Cys Pro Gln Val
1365 1370 1375

Leu Ile His Lys His Leu Ala Arg Leu Leu Ser Val Ile Leu Pro Asn
1380 1385 1390

Leu Gln Ser Glu Asn Thr Pro Gly Leu Leu Ser Val Asp Leu Phe His
1395 1400 1405

1075 1080 1085 1090 1095 1100 1105 1110 1115 1120 1125 1130 1135 1140 1145 1150 1155 1160 1165 1170 1175 1180 1185 1190 1195 1200 1205 1210 1215 1220 1225 1230 1235 1240 1245 1250 1255 1260 1265 1270 1275 1280 1285 1290 1295 1300 1305 1310 1315 1320 1325 1330 1335 1340 1345 1350 1355 1360 1365 1370 1375 1380 1385 1390 1395 1400 1405

Val Leu Val Gly Ala Val Leu Ala Phe Pro Ser Leu Tyr Trp Asp Asp
1410 1415 1420

Thr Val Asp Leu Gln Pro Ser Pro Leu Ser Ser Ser Tyr Asn His Leu
1425 1430 1435 1440

Tyr Leu Phe His Leu Ile Thr Met Ala His Met Leu Gln Ile Leu Leu
1445 1450 1455

Thr Thr Asp Thr Asp Leu Ser Pro Gly Pro Pro Leu Ala Glu Gly Glu
1460 1465 1470

Glu Asp Ser Glu Glu Ala Arg Cys Ala Ser Ala Phe Phe Val Glu Val
1475 1480 1485

Ser Gln His Thr Asp Gly Leu Thr Gly Cys Gly Ala Pro Gly Trp Tyr
1490 1495 1500

Leu Trp Leu Ser Leu Arg Asn Gly Ile Thr Pro Tyr Leu Arg Cys Ala
1505 1510 1515 1520

Ala Leu Leu Phe His Tyr Leu Leu Gly Val Ala Pro Pro Glu Glu Leu
1525 1530 1535

Phe Ala Asn Ser Ala Glu Gly Glu Phe Ser Ala Leu Cys Ser Tyr Leu
1540 1545 1550

Ser Leu Pro Thr Asn Leu Phe Leu Leu Phe Gln Glu Tyr Trp Asp Thr
1555 1560 1565

Ile Arg Pro Leu Leu Gln Arg Trp Cys Gly Asp Pro Ala Leu Leu Lys
1570 1575 1580

Ser Leu Lys Gln Lys Ser Ala Val Val Arg Tyr Pro Arg Lys Arg Asn
1585 1590 1595 1600

Ser Leu Ile Glu Leu Pro Glu Asp Tyr Ser Cys Leu Leu Asn Gln Ala
1605 1610 1615

Ser His Phe Arg Cys Pro Arg Ser Ala Asp Asp Glu Arg Lys His Pro
1620 1625 1630

Val Leu Cys Leu Phe Cys Gly Ala Ile Leu Cys Ser Gln Asn Ile Cys
1635 1640 1645

Cys Gln Glu Ile Val Asn Gly Glu Glu Val Gly Ala Cys Val Phe His
1650 1655 1660

Ala Leu His Cys Gly Ala Gly Val Cys Ile Phe Leu Lys Ile Arg Glu
1665 1670 1675 1680

Cys Arg Val Val Leu Val Glu Gly Lys Ala Arg Gly Cys Ala Tyr Pro
1685 1690 1695

Ala Pro Tyr Leu Asp Glu Tyr Gly Glu Thr Asp Pro Gly Leu Lys Arg
1700 1705 1710

Gly Asn Pro Leu His Leu Ser Arg Glu Arg Tyr Arg Lys Leu His Leu
1715 1720 1725

Val Trp Gln Gln His Cys Ile Ile Glu Glu Ile Ala Arg Ser Gln Glu
1730 1735 1740

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Thr Asn Gln Met Leu Phe Gly Phe Asn Trp Gln Leu Leu
1745 1750 1755

<210> 16
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: peptide

<400> 16
Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg
1 5 10

<210> 17
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: peptide

<400> 17
Gly Gly Gly Gly Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg
1 5 10 15

<210> 18
<211> 5205
<212> DNA
<213> Homo sapiens

<400> 18
atggcggacg aggaggctgg aggtactgag aggatggaaa tcagcgcgga gttaccccag 60
accctcagc gtctggcatc ttggtgggat cagcaagttg atttttatac tgctttcttg 120
catcatttgg cacaattggt gccagaaatt tactttgctg aaatggaccc agacttggaa 180
aagcaggagg aaagtgtaca aatgtcaata ttactccac tggaatggta cttatttggga 240
gaagatccag atatttgctt agagaaattg aagcacagtg gagcatttca gctttgtggg 300
agggttttca aaagtggaga gacaacctat tcttgcaggg attgtgcaat tgatccaaca 360
tgtgtactct gtatggactg cttccaggac agtgttcata aaaatcatcg ttacaagatg 420
catacttcta ctggaggagg gttctgtgac tgtggagaca cagaggcatg gaaaactggc 480
cctttttgtg taaatcatga acctggaaga gcaggacta taaaagagaa ttcacgctgt 540
ccgttgaatg aagaggtaat tgtccaagcc aggaaaatat ttccttcagt gataaaatat 600
gtcgtagaaa tgactatatg ggaagaggaa aaagaactgc ctctgaact ccagataagg 660
knryycvndh hsydhgtcat atacagccta caaagagctc ttgactgtga gctcgcagag 720
gccagttgc ataccactgc cattgacaaa gaggtcgtc gggctgttaa agcgggagct 780
tatgctgctt gccaggaagc aaaggaagat ataaagagtc attcagaaaa tgtctctcaa 840

catccacttc atgtagaagt attacactca gagattatgg ctcatcagaa atttgctttg 900
cgtcttggtt cctggatgaa caaaattatg agctattcaa gtgacttttag gcagatcttt 960
tgccaagcat gccttagaga agaacctgac tcggagaatc cctgtctcat aagcaggtta 1020
atgctttggg atgcaaagct ttataaaggt gcccgtaaga tccttcatga attgatcttc 1080
agcagttttt ttatggagat ggaatacaaa aaactctttg ctatggaatt tgtgaagtat 1140
tataaacaac tgcagaaaga atatatcagt gatgatcatg acagaagtat ctctataact 1200
gcactttcag ttcagatgtt tactgttcct actctggctc gacatcttat tgaagagcag 1260
aatgttatct ctgtcattac tgaaactctg ctagaagttt tacctgagta cttggacagg 1320
aacaataaat tcaacttcca gggttatagc caggacaaat tgggaagagt atatgcagta 1380
atatgtgacc taaagtatat cctgatcagc aaaccacaaa tatggacaga aagattaaga 1440
atgcagttcc ttgaaggttt tcgatctttt ttgaagattc ttacctgtat gcagggaatg 1500
gaagaaatcc gaagacaggt tgggcaacac attgaagtgg atcctgattg ggaggctgcc 1560
attgctatac agatgcaatt gaagaatatt ttactcatgt tccaagagtg gtgtgcttgt 1620
gatgaagaac tcttacttgt ggcttataaa gaatgtcaca aagctgtgat gaggtgcagt 1680
accagtttca tatctagtag caagacagta gtacaatcgt gtggacatag tttggaaaca 1740
aagtcctaca gagtatctga ggatcttgta agcatacatc tgccactctc taggaccctt 1800
gctggctctc atgtacgttt aagcaggctg ggtgctgttt caagactgca tgaatttgtg 1860
tcttttgagg actttcaagt agaggtaacta gtggaatata ctttacgttg tctggtgttg 1920
gttgcccagg ttgttgctga gatgtggcga agaaatggac tgtctcttat tagccagggtg 1980
ttttattacc aagatgttaa gtgcagagaa gaaatgtatg ataaagatat catcatgctt 2040
cagattggtg catctttaat ggatcccaat aagttcttgt tactggtact tcagaggtat 2100
gaacttgccg aggcttttaa caagaccata tctacaaaag accaggattt gattaaacaa 2160
tataatacac taatagaaga aatgcttcag gtccctcatct atattgtggg tgagcgttat 2220
gtacctggag tgggaaatgt gaccaaagaa gaggtcacao tgagagaaat cattcacttg 2280
ctttgcattg aacctatgcc acacagtgcc attgccaaaa atttacctga gaatgaaaat 2340
aatgaaactg gcttagagaa tgtcataaac aaagtggcca catttaagaa accagggtga 2400
tcaggccatg gagtttatga actaaaagat gaatcactga aagacttcaa tatgtacttt 2460
tatcattact ccaaaaccca gcatagcaag gctgaacata tgcagaagaa aaggagaaaa 2520
caagaaaaca aagatgaagc attgccgcca ccaccacctc ctgaattctg ccctgctttc 2580
agcaaagtga ttaaccttct caactgtgat atcatgatgt acattctcag gaccgtattt 2640
gagcgggcaa tagacacaga ttctaacttg tggaccgaag ggatgctcca aatggctttt 2700
catattctgg cattgggttt actagaagag aagcaacagc ttcaaaaagc tcctgaagaa 2760

gaagtaacat ttgacttttta tcataaggct tcaagattgg gaagttcagc catgaatata 2820
caaagtcttt tggaaaaact caaaggaatt cccagtttag aaggccagaa ggacatgata 2880
acgtggatac ttcagatggt tgacacagtg aagcgattaa gagaaaaatc ttgtttaatt 2940
gtagcaacca catcaggatc ggaatctatt aagaatgatg agattactca tgataaagaa 3000
aaagcagaac gaaaaagaaa agctgaagct gctaggctac atcgccagaa gatcatggct 3060
cagatgtctg ccttacagaa aaacttcatt gaaactcata aactcatgta tgacaatata 3120
tcagaaatgc ctgggaaaga agattccatt atggaggaag agagcacccc agcagtcagt 3180
gactactcta gaattgcttt gggctctaaa cggggtccat ctgttactga aaaggaggtg 3240
ctgacgtgca tcctttgcc aagaagacag gaggtgaaaa tagaaaataa tgccatggta 3300
ttatcggcct gtgtccagaa atctactgcc ttaaccagc acaggggaaa acccatagaa 3360
ctctcaggag aagccctaga cccacttttc atggatccag acttggcata tggaacttat 3420
acaggaagct gtggtcatgt aatgcacgca gtgtgctggc agaagtattt tgaagctgta 3480
cagctgagct ctcagcagcg cattcatggt gacctttttg acttggaaag tggagaatat 3540
ctttgccctc tttgcaaata tctgtgcaat actgtgatcc ccattattcc tttgcaacct 3600
caaaagataa acagtgagaa tgcagatgct cttgctcaac ttttgacctt ggcacggtgg 3660
atacagactg ttctggccag aatatcaggt tataatataa gacatgctaa aggagaaaac 3720
ccaattccta ttttctttaa tcaaggaatg ggagattcta ctttggagtt ccattccatc 3780
ctgagttttg gcgttgagtc ttcgattaaa tattcaaata gcatcaagga aatgggtatt 3840
ctctttgcc caacaattta tagaattgga ttgaaagtgc cacctgatga aagggatcct 3900
cgagtcccca tgctgacctg gagcacctgc gctttcacta tccaggcaat tgaaaatcta 3960
ttgggagatg aagggaaacc tctgtttgga gcacttcaaa ataggcagca taatgggtctg 4020
aaagcattaa tgcagtttgc agttgcacag aggattacct gtcctcaggt cctgatacag 4080
aaacatctgg ttcgtcttct atcagttggt cttcctaaca taaaatcaga agatacacca 4140
tgcccttctgt ctatagatct gtttcatggt ttgggtgggtg ctgtgttagc attcccatcc 4200
ttgtattggg atgaccctgt tgatctgcag ccttcttcag ttagttcttc ctataaccac 4260
ctttatctct tccatttgat caccatggca cacatgcttc agatactact tacagtagac 4320
acaggcctac cccttgctca ggttcaagaa gacagtgaag aggcctatc cgcactctct 4380
ttctttgcag aaatttctca atatacaagt ggctccattg ggtgtgatat tcctggctgg 4440
tatttggtgg tctcactgaa gaatggcatc accccttata ttcgtgtgct tgcattgttt 4500
ttccactatt tacttggggg aactccgcct gaggaactgc ataccaatc tgcagaagga 4560
gagtacagtg cactctgtag ctatctatct ttacctaaa atttgttctt gctcttccag 4620
gaatattggg atactgtaag gcccttgctc cagaggcggg gtgcagatcc tgccttacta 4680

aactgtttga agcaaaaaaa caccgtgggc aggtacccta gaaaaagaaa tagtttgata 4740
gagcttctctg atgactatag ctgcctcctg aatcaagctt ctcatttcag gtgcccacgg 4800
tctgcagatg atgagcgaaa gcacccctgtc ctctgccttt tctgtggggc tatactatgt 4860
tctcagaaca tttgttgcca ggaaattgtg aacggggaag aggttggagc ttgcattttt 4920
cacgcacttc actgtggagc cggagtctgc attttcctaa aaatcagaga atgccgagtg 4980
gtcctgggtg aaggtaaagc cagaggctgt gcctatccag ctccttactt ggatgaatat 5040
ggagaaacag accctggcct gaagaggggc aaccccttc atttatctcg tgagcgggtat 5100
cggaagctcc atttgggtctg gcaacaacac tgcattatag aagagattgc taggagccaa 5160
gagactaatc agatgttatt tggattcaac tggcagttac tgtga 5205

<210> 19
<211> 1735
<212> PRT
<213> Homo sapiens

<400> 19
Ala Met Glu Gly Asn Met Ala Asp Glu Glu Ala Gly Gly Thr Glu Arg
1 5 10 15
Met Glu Ile Ser Ala Glu Leu Pro Gln Thr Pro Gln Arg Leu Ala Ser
20 25 30
Trp Trp Asp Gln Gln Val Asp Phe Tyr Thr Ala Phe Leu His His Leu
35 40 45
Ala Gln Leu Val Pro Glu Ile Tyr Phe Ala Glu Met Asp Pro Asp Leu
50 55 60
Glu Lys Gln Glu Glu Ser Val Gln Met Ser Ile Phe Thr Pro Leu Glu
65 70 75 80
Trp Tyr Leu Phe Gly Glu Asp Pro Asp Ile Cys Leu Glu Lys Leu Lys
85 90 95
His Ser Gly Ala Phe Gln Leu Cys Gly Arg Val Phe Lys Ser Gly Glu
100 105 110
Thr Thr Tyr Ser Cys Arg Asp Cys Ala Ile Asp Pro Thr Cys Val Leu
115 120 125
Cys Met Asp Cys Phe Gln Asp Ser Val His Lys Asn His Arg Tyr Lys
130 135 140
Met His Thr Ser Thr Gly Gly Gly Phe Cys Asp Cys Gly Asp Thr Glu
145 150 155 160
Ala Trp Lys Thr Gly Pro Phe Cys Val Asn His Glu Pro Gly Arg Ala
165 170 175
Gly Thr Ile Lys Glu Asn Ser Arg Cys Pro Leu Asn Glu Glu Val Ile
180 185 190
Val Gln Ala Arg Lys Ile Phe Pro Ser Val Ile Lys Tyr Val Val Glu
195 200 205

Met	Thr	Ile	Trp	Glu	Glu	Glu	Lys	Glu	Leu	Pro	Pro	Glu	Leu	Gln	Ile
210							215				220				
Arg	Glu	Lys	Asn	Glu	Arg	Tyr	Tyr	Cys	Val	Leu	Phe	Asn	Asp	Glu	His
225					230					235					240
His	Ser	Tyr	Asp	His	Val	Ile	Tyr	Ser	Leu	Gln	Arg	Ala	Leu	Asp	Cys
				245					250					255	
Glu	Leu	Ala	Glu	Ala	Gln	Leu	His	Thr	Thr	Ala	Ile	Asp	Lys	Glu	Gly
			260					265					270		
Arg	Arg	Ala	Val	Lys	Ala	Gly	Ala	Tyr	Ala	Ala	Cys	Gln	Glu	Ala	Lys
		275					280					285			
Glu	Asp	Ile	Lys	Ser	His	Ser	Glu	Asn	Val	Ser	Gln	His	Pro	Leu	His
290						295					300				
Val	Glu	Val	Leu	His	Ser	Glu	Ile	Met	Ala	His	Gln	Lys	Phe	Ala	Leu
305					310					315					320
Arg	Leu	Gly	Ser	Trp	Met	Asn	Lys	Ile	Met	Ser	Tyr	Ser	Ser	Asp	Phe
				325					330					335	
Arg	Gln	Ile	Phe	Cys	Gln	Ala	Cys	Leu	Arg	Glu	Glu	Pro	Asp	Ser	Glu
			340					345					350		
Asn	Pro	Cys	Leu	Ile	Ser	Arg	Leu	Met	Leu	Trp	Asp	Ala	Lys	Leu	Tyr
		355					360					365			
Lys	Gly	Ala	Arg	Lys	Ile	Leu	His	Glu	Leu	Ile	Phe	Ser	Ser	Phe	Phe
	370					375					380				
Met	Glu	Met	Glu	Tyr	Lys	Lys	Leu	Phe	Ala	Met	Glu	Phe	Val	Lys	Tyr
385					390					395					400
Tyr	Lys	Gln	Leu	Gln	Lys	Glu	Tyr	Ile	Ser	Asp	Asp	His	Asp	Arg	Ser
				405					410					415	
Ile	Ser	Ile	Thr	Ala	Leu	Ser	Val	Gln	Met	Phe	Thr	Val	Pro	Thr	Leu
			420					425					430		
Ala	Arg	His	Leu	Ile	Glu	Glu	Gln	Asn	Val	Ile	Ser	Val	Ile	Thr	Glu
		435					440					445			
Thr	Leu	Leu	Glu	Val	Leu	Pro	Glu	Tyr	Leu	Asp	Arg	Asn	Asn	Lys	Phe
						455					460				
Asn	Phe	Gln	Gly	Tyr	Ser	Gln	Asp	Lys	Leu	Gly	Arg	Val	Tyr	Ala	Val
465					470					475					480
Ile	Cys	Asp	Leu	Lys	Tyr	Ile	Leu	Ile	Ser	Lys	Pro	Thr	Ile	Trp	Thr
				485					490					495	
Glu	Arg	Leu	Arg	Met	Gln	Phe	Leu	Glu	Gly	Phe	Arg	Ser	Phe	Leu	Lys
			500					505					510		
Ile	Leu	Thr	Cys	Met	Gln	Gly	Met	Glu	Glu	Ile	Arg	Arg	Gln	Val	Gly
		515					520					525			
Gln	His	Ile	Glu	Val	Asp	Pro	Asp	Trp	Glu	Ala	Ala	Ile	Ala	Ile	Gln
						535						540			

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

Met 545	Gln	Leu	Lys	Asn	Ile 550	Leu	Leu	Met	Phe	Gln	Glu	Trp	Cys	Ala	Cys 560
Asp	Glu	Glu	Leu	Leu 565	Leu	Val	Ala	Tyr	Lys 570	Glu	Cys	His	Lys	Ala 575	Val
Met	Arg	Cys	Ser 580	Thr	Ser	Phe	Ile	Ser 585	Ser	Ser	Lys	Thr	Val 590	Val	Gln
Ser	Cys	Gly 595	His	Ser	Leu	Glu	Thr 600	Lys	Ser	Tyr	Arg	Val 605	Ser	Glu	Asp
Leu 610	Val	Ser	Ile	His	Leu	Pro 615	Leu	Ser	Arg	Thr 620	Leu	Ala	Gly	Leu	His
Val 625	Arg	Leu	Ser	Arg	Leu	Gly 630	Ala	Val	Ser	Arg 635	Leu	His	Glu	Phe	Val 640
Ser	Phe	Glu	Asp 645	Phe	Gln	Val	Glu	Val 650	Leu	Val	Glu	Tyr	Pro 655	Leu	Arg
Cys	Leu	Val	Leu 660	Val	Ala	Gln	Val 665	Val	Ala	Glu	Met	Trp	Arg 670	Arg	Asn
Gly	Leu 675	Ser	Leu	Ile	Ser	Gln	Val 680	Phe	Tyr	Tyr	Gln 685	Asp	Val	Lys	Cys
Arg 690	Glu	Glu	Met	Tyr	Asp	Lys 695	Asp	Ile	Ile	Met	Leu 700	Gln	Ile	Gly	Ala
Ser 705	Leu	Met	Asp	Pro 710	Asn	Lys	Phe	Leu	Leu 715	Leu	Val	Leu	Gln	Arg	Tyr 720
Glu	Leu	Ala	Glu 725	Ala	Phe	Asn	Lys	Thr 730	Ile	Ser	Thr	Lys	Asp	Gln 735	Asp
Leu	Ile	Lys 740	Gln	Tyr	Asn	Thr	Leu 745	Ile	Glu	Glu	Met	Leu 750	Gln	Val	Leu
Ile	Tyr 755	Ile	Val	Gly	Glu	Arg	Tyr 760	Val	Pro	Gly	Val 765	Gly	Asn	Val	Thr
Lys 770	Glu	Glu	Val	Thr	Met 775	Arg	Glu	Ile	Ile	His 780	Leu	Leu	Cys	Ile	Glu
Pro 785	Met	Pro	His	Ser 790	Ala	Ile	Ala	Lys	Asn 795	Leu	Pro	Glu	Asn	Glu	Asn 800
Asn	Glu	Thr	Gly 805	Leu	Glu	Asn	Val	Ile 810	Asn	Lys	Val	Ala	Thr 815	Phe	Lys
Lys	Pro	Gly 820	Val	Ser	Gly	His	Gly 825	Val	Tyr	Glu	Leu	Lys 830	Asp	Glu	Ser
Leu	Lys 835	Asp	Phe	Asn	Met	Tyr	Phe 840	Tyr	His	Tyr	Ser 845	Lys	Thr	Gln	His
Ser 850	Lys	Ala	Glu	His	Met	Gln 855	Lys	Lys	Arg	Arg	Lys 860	Gln	Glu	Asn	Lys
Asp 865	Glu	Ala	Leu	Pro 870	Pro	Pro	Pro	Pro	Pro 875	Glu	Phe	Cys	Pro	Ala	Phe 880

Ser Lys Val Ile Asn Leu Leu Asn Cys Asp Ile Met Met Tyr Ile Leu
885 890 895

Arg Thr Val Phe Glu Arg Ala Ile Asp Thr Asp Ser Asn Leu Trp Thr
900 905 910

Glu Gly Met Leu Gln Met Ala Phe His Ile Leu Ala Leu Gly Leu Leu
915 920 925

Glu Glu Lys Gln Gln Leu Gln Lys Ala Pro Glu Glu Glu Val Thr Phe
930 935 940

Asp Phe Tyr His Lys Ala Ser Arg Leu Gly Ser Ser Ala Met Asn Ile
945 950 955 960

Gln Met Leu Leu Glu Lys Leu Lys Gly Ile Pro Gln Leu Glu Gly Gln
965 970 975

Lys Asp Met Ile Thr Trp Ile Leu Gln Met Phe Asp Thr Val Lys Arg
980 985 990

Leu Arg Glu Lys Ser Cys Leu Ile Val Ala Thr Thr Ser Gly Ser Glu
995 1000 1005

Ser Ile Lys Asn Asp Glu Ile Thr His Asp Lys Glu Lys Ala Glu Arg
1010 1015 1020

Lys Arg Lys Ala Glu Ala Ala Arg Leu His Arg Gln Lys Ile Met Ala
1025 1030 1035 1040

Gln Met Ser Ala Leu Gln Lys Asn Phe Ile Glu Thr His Lys Leu Met
1045 1050 1055

Tyr Asp Asn Thr Ser Glu Met Pro Gly Lys Glu Asp Ser Ile Met Glu
1060 1065 1070

Glu Glu Ser Thr Pro Ala Val Ser Asp Tyr Ser Arg Ile Ala Leu Gly
1075 1080 1085

Pro Lys Arg Gly Pro Ser Val Thr Glu Lys Glu Val Leu Thr Cys Ile
1090 1095 1100

Leu Cys Gln Glu Glu Gln Glu Val Lys Ile Glu Asn Asn Ala Met Val
1105 1110 1115 1120

Leu Ser Ala Cys Val Gln Lys Ser Thr Ala Leu Thr Gln His Arg Gly
1125 1130 1135

Lys Pro Ile Glu Leu Ser Gly Glu Ala Leu Asp Pro Leu Phe Met Asp
1140 1145 1150

Pro Asp Leu Ala Tyr Gly Thr Tyr Thr Gly Ser Cys Gly His Val Met
1155 1160 1165

His Ala Val Cys Trp Gln Lys Tyr Phe Glu Ala Val Gln Leu Ser Ser
1170 1175 1180

Gln Gln Arg Ile His Val Asp Leu Phe Asp Leu Glu Ser Gly Glu Tyr
1185 1190 1195 1200

Leu Cys Pro Leu Cys Lys Ser Leu Cys Asn Thr Val Ile Pro Ile Ile
1205 1210 1215

Pro Leu Gln Pro Gln Lys Ile Asn Ser Glu Asn Ala Asp Ala Leu Ala
1220 1225 1230

Gln Leu Leu Thr Leu Ala Arg Trp Ile Gln Thr Val Leu Ala Arg Ile
1235 1240 1245

Ser Gly Tyr Asn Ile Arg His Ala Lys Gly Glu Asn Pro Ile Pro Ile
1250 1255 1260

Phe Phe Asn Gln Gly Met Gly Asp Ser Thr Leu Glu Phe His Ser Ile
1265 1270 1275 1280

Leu Ser Phe Gly Val Glu Ser Ser Ile Lys Tyr Ser Asn Ser Ile Lys
1285 1290 1295

Glu Met Val Ile Leu Phe Ala Thr Thr Ile Tyr Arg Ile Gly Leu Lys
1300 1305 1310

Val Pro Pro Asp Glu Arg Asp Pro Arg Val Pro Met Leu Thr Trp Ser
1315 1320 1325

Thr Cys Ala Phe Thr Ile Gln Ala Ile Glu Asn Leu Leu Gly Asp Glu
1330 1335 1340

Gly Lys Pro Leu Phe Gly Ala Leu Gln Asn Arg Gln His Asn Gly Leu
1345 1350 1355 1360

Lys Ala Leu Met Gln Phe Ala Val Ala Gln Arg Ile Thr Cys Pro Gln
1365 1370 1375

Val Leu Ile Gln Lys His Leu Val Arg Leu Leu Ser Val Val Leu Pro
1380 1385 1390

Asn Ile Lys Ser Glu Asp Thr Pro Cys Leu Leu Ser Ile Asp Leu Phe
1395 1400 1405

His Val Leu Val Gly Ala Val Leu Ala Phe Pro Ser Leu Tyr Trp Asp
1410 1415 1420

Asp Pro Val Asp Leu Gln Pro Ser Ser Val Ser Ser Ser Tyr Asn His
1425 1430 1435 1440

Leu Tyr Leu Phe His Leu Ile Thr Met Ala His Met Leu Gln Ile Leu
1445 1450 1455

Leu Thr Val Asp Thr Gly Leu Pro Leu Ala Gln Val Gln Glu Asp Ser
1460 1465 1470

Glu Glu Ala His Ser Ala Ser Ser Phe Phe Ala Glu Ile Ser Gln Tyr
1475 1480 1485

Thr Ser Gly Ser Ile Gly Cys Asp Ile Pro Gly Trp Tyr Leu Trp Val
1490 1495 1500

Ser Leu Lys Asn Gly Ile Thr Pro Tyr Leu Arg Cys Ala Ala Leu Phe
1505 1510 1515 1520

Phe His Tyr Leu Leu Gly Val Thr Pro Pro Glu Glu Leu His Thr Asn
1525 1530 1535

Ser Ala Glu Gly Glu Tyr Ser Ala Leu Cys Ser Tyr Leu Ser Leu Pro
1540 1545 1550

Protein Data Bank

Thr Asn Leu Phe Leu Leu Phe Gln Glu Tyr Trp Asp Thr Val Arg Pro
1555 1560 1565

Leu Leu Gln Arg Arg Cys Ala Asp Pro Ala Leu Leu Asn Cys Leu Lys
1570 1575 1580

Gln Lys Asn Thr Val Val Arg Tyr Pro Arg Lys Arg Asn Ser Leu Ile
1585 1590 1595 1600

Glu Leu Pro Asp Asp Tyr Ser Cys Leu Leu Asn Gln Ala Ser His Phe
1605 1610 1615

Arg Cys Pro Arg Ser Ala Asp Asp Glu Arg Lys His Pro Val Leu Cys
1620 1625 1630

Leu Phe Cys Gly Ala Ile Leu Cys Ser Gln Asn Ile Cys Cys Gln Glu
1635 1640 1645

Ile Val Asn Gly Glu Glu Val Gly Ala Cys Ile Phe His Ala Leu His
1650 1655 1660

Cys Lys Ala Arg Gly Cys Ala Tyr Pro Ala Pro Tyr Leu Asp Glu Tyr
1665 1670 1675 1680

Gly Glu Thr Asp Pro Gly Leu Lys Arg Gly Asn Pro Leu His Leu Ser
1685 1690 1695

Arg Glu Arg Tyr Arg Lys Leu His Leu Val Trp Gln Gln His Cys Ile
1700 1705 1710

Ile Glu Glu Ile Ala Arg Ser Gln Glu Thr Asn Gln Met Leu Phe Gly
1715 1720 1725

Phe Asn Trp Gln Leu Leu *
1730 1735

<210> 20
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR Primer

<400> 20
agaaggagag tacagtgcac tc

22

<210> 21
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR Primer

<400> 21
cgaaagcatc ctgtcctctg

20

<210> 22
<211> 18

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR Primer

<400> 22
aggaagctgt ggtcatgt 18

<210> 23
<211> 14
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR Primer

<400> 23
gttaggaaga actg 14

<210> 24
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR Primer

<400> 24
aagaacagcg aaggcaacag 20

<210> 25
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR Primer

<400> 25
cgcagctacc ccaacacatt ct 22

<210> 26
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR Primer

<400> 26
tttcttccat tccctgcata ca 22

<210> 27
<211> 25
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PCR Primer

<400> 27

caaaacttta taaaggtgcc cgtaa

25

<210> 28

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PCR Primer

<400> 28

attccctgca tgcacttcag taa

23

<210> 29

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PCR Primer

<400> 29

cattccctgc atgcacttca g

21